MODELING TRAFFIC-ACTUATED CONTROL WITH TRANSYT-7F

By

DAVID HALE

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2000

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MODELING TRAFFIC-ACTUATED CONTROL WITH TRANSYT-7F

By

David Hale

August 2000

Chair: Kenneth G. Courage
Major Department: Civil Engineering

Research suggests that traffic congestion in U.S. cities has grown rapidly in recent years, and that numerous solutions are needed to address the problem. This dissertation describes new research for producing basic improvements to the practice of traffic signal timing optimization, in order to improve one of the many available weapons for fighting congestion and delay.

TRANSYT-7F is one of the most comprehensive signal timing tools in existence, and has evolved into a benchmark within the transportation profession. Although TRANSYT-7F was developed in an era of pre-timed signal control, it was modified in the 1980s to automatically estimate the average green times for actuated controllers.

Although effective at eliminating wasted green time, the existing actuated control sub-model within TRANSYT-7F is oversimplified. The model does not recognize most of the signal settings associated with today's controllers. In addition, this and other actuated control models from the literature do not recognize numerous operational characteristics affecting phase times. The end result is that performance estimates and timing plans generated by the program are potentially less accurate and optimal.

The overall goal of this study was to develop an improved actuated control methodology for usage within TRANSYT-7F, and perhaps within other programs and procedures. A literature review was performed to ascertain available technology on actuated control including phase time calculation, vehicle delay estimation, and control parameter design or optimization. Subsequently, certain models from the literature were chosen as candidates for improvement of TRANSYT-7F performance. In addition, new prototype models were developed during the course of the study. Along with existing literature models, these prototypes were subjected to a battery of tests in order to scrutinize their strengths and weaknesses. The well-known CORSIM simulation program was used as a reference point for model comparison.

Significant amounts of experimental data revealed dramatic improvement in the accuracy of actuated phase time calculation when the candidate models were applied by TRANSYT-7F. A smaller amount of data was also collected to demonstrate improvements in the optimal timing plans, based on the new methodology. Conclusions and recommendations, regarding new modeling capabilities and future research, are provided in the final chapter.

CHAPTER 1
INTRODUCTION

Research from the Texas Transportation Institute (TTI) suggests that traffic congestion in U.S. cities has grown rapidly in recent years, and that "a full array of solutions and measures are essential in addressing the mobility problem" [Lomax and Schrank, 1998]. Simple solutions such as carpooling, or building additional roadway lanes, are reportedly inadequate for dealing with increasing traffic problems. Lomax and Schrank report that in 1997, the financial cost of congestion exceeded $72 billion per year, up from $66 billion in 1996, and that increases in delay have been more prevalent in small- to medium-sized cities than in the nation's largest cities.

Numerous congestion-reducing strategies have been proposed for improving transportation efficiency in an existing roadway network. Some strategies involve modification of driver behavior (e.g., the use of public transit), whereas other strategies involve better traffic management. This dissertation focuses on one aspect of better traffic management, namely the optimization of traffic signal timing plans in centrally coordinated traffic control systems. While no single device or strategy can be expected to "solve" the problem of traffic congestion by itself, signal timing optimization has been

demonstrated to be an extremely cost-effective strategy. When properly applied, it has proven effective in reducing delay, stops, fuel consumption and other measures related to operating costs and driver-perceived disutility [Deakin et al., 1984]. Any technique that offers a potential improvement in traffic control system performance should provide a useful contribution in combating this growing national problem.

Problem Statement

The importance of efficient traffic signal operation has been recognized for many years. Various signal system simulation and optimization models are now available for analysts and engineers. Simulation models offer a realistic interpretation of traffic flow and performance, but their function is limited to evaluating the performance of a specified operational alternative. In other words, they do not provide features that explicitly optimize the performance. Optimization models, as the name implies, do attempt to find the "best" set of operating parameters for a specified performance objective. However, they often lack certain detailed treatment of traffic flow characteristics found within simulation models. The simplifying assumptions required to support a productive optimization process have, in many cases, compromised the quality of the final product. The result is that the timing plans being implemented in many traffic control systems today still have significant room for improvement.

One of the most promising areas for improvement is the modeling of "traffic-actuated" control, in which instantaneous information from traffic detectors is incorporated into the control tactics for fine-tuning intersection performance. Existing

optimization models apply the simplifying assumption of "pre-timed" control, in which the operation at each intersection is characterized by a series of fixed-duration intervals. The optimization process determines the optimal duration and position of each interval. Some adjustments are made in an attempt to represent the effects of traffic-actuated control, but the results to this point have not proven to be entirely satisfactory.

Objectives

The overall goal of this study is to develop an improved treatment of traffic-actuated control for application within deterministic optimization models. Regarding the type of actuated control to be analyzed, the scope of this study involves "basic" actuated control, where presence detectors are installed at the stop line. To accomplish the overall goal, it is necessary at the outset to choose a specific optimization model for improvement, as well as a simulation model for evaluation of the improvements. These two choices will establish the scope of the project.

The Traffic Network Study Tool (TRANSYT) program [Robertson, 1968] is the most logical choice for the optimization model. TRANSYT was developed by the Transport Research Laboratory in the United Kingdom. It is used extensively throughout the world for traffic control system timing design and evaluation. The specific version of the TRANSYT model to be used in this study will be TRANSYT-7F, release 8. TRANSYT-7F was developed as a derivative work by the University of Florida Transportation Research Center [Wallace et al., 1981]. It is distributed worldwide by the University of Florida and currently has approximately 1,400 registered agencies

worldwide. The extensive recognition of TRANSYT-7F, combined with the institutional support of the University of Florida, make it a natural choice for purposes of this project. More detailed descriptions of the TRANSYT-7F modeling and calibration process are presented in the appendix.

Ideally testing and validation of the project results would be accomplished through the collection and analysis of field data; however, an adequate empirical validation would require resources several orders of magnitude beyond those available to this project. This limitation dictates the need for simulation as a surrogate for field data collection.

The choice of a simulation model is not difficult in this case. The CORridor SIMulation (CORSIM) [ITT Systems and Sciences Corporation, 1998] model has been developed and enhanced continually over the past 25 years by the Federal Highway Administration (FHWA). It was designed specifically for the purpose intended by this project (i.e., a surrogate for field data) and is used extensively for this purpose.

CORSIM is especially strong in the treatment of traffic-actuated control. It models this type of control explicitly, using a software module developed from the same source code as the real-time control logic contained within the actual field hardware. Chapter 2 (literature review) contains a reference that illustrates field validation of CORSIM's actuated phase times. Therefore CORSIM will be used for testing and validation of the optimization enhancements to be developed as a part of this project.

Consider, for example, the comparison between signal phase times estimated by a deterministic optimization model (TRANSYT-7F) and a detailed microscopic simulation model (CORSIM), as shown in figure 1-1.

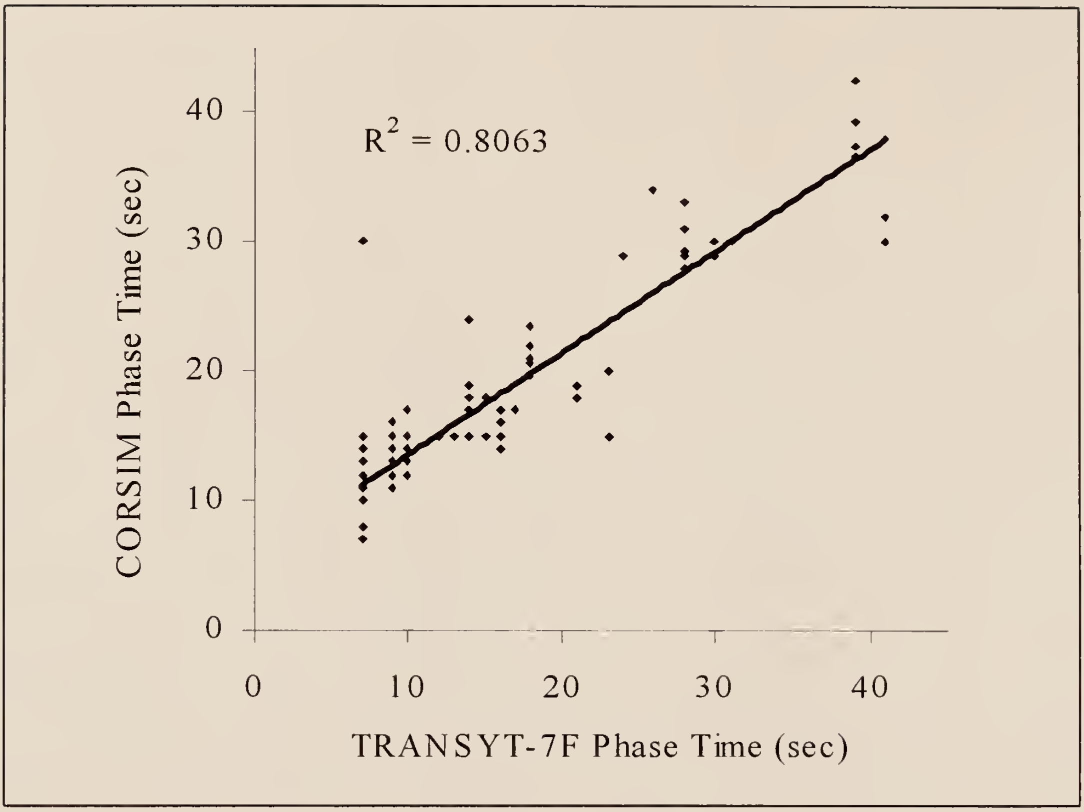


Figure 1-1: Actuated Phase Time Comparison – CORSIM vs. TRANSYT-7F

The level of correlation evident in this comparison suggests that the optimization model is not replicating simulation model conditions satisfactorily. A corollary to this observation is that the "optimal" design (i.e., the final product of the optimization process) will be based on incorrect information and will not therefore realize its potential for optimizing system performance. It is therefore possible, and definitely desirable, to improve the treatment of traffic-actuated control in the design and optimization of traffic control system timing plans. This would provide the opportunity to improve efficiency of performance at thousands of signalized intersections throughout the USA.

As an alternative, consider an updated comparison between phase times estimated by TRANSYT-7F and CORSIM, given the new methodologies to be presented within the body of this study, as shown in figure 1-2.

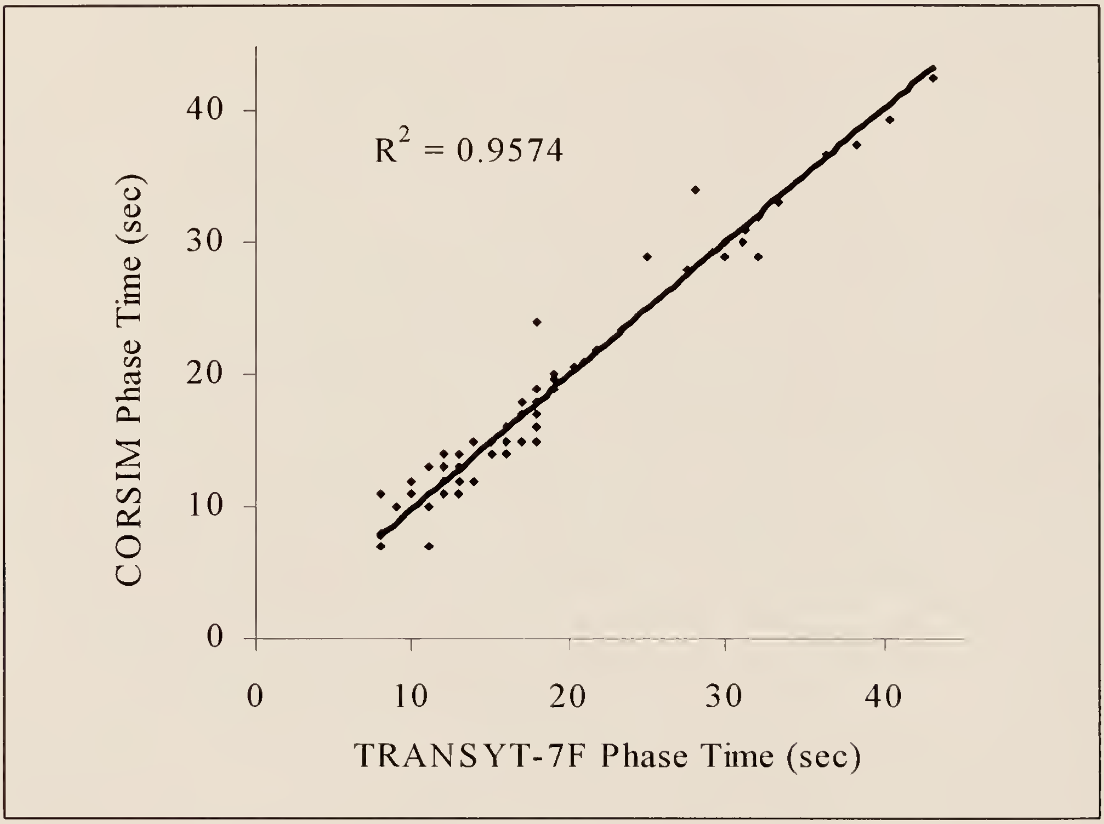


Figure 1-2: Updated Phase Time Comparison – CORSIM vs. TRANSYT-7F

TRANSYT-7F and other optimization models are based on a series of simulation or evaluation runs. Typically the simulation or evaluation run that results in the best performance is reported as optimal. In other words, the effectiveness of optimization is predicated on the accuracy of simulation or evaluation. Because of this, improvements in actuated phase time calculation (as illustrated in figures 1-1 and 1-2) can result in an improved optimization process. Table 1-1 shows an example of traffic network

performance improvements that are possible under the updated methodologies for actuated control.

Table 1-1: Sample Optimization Based on Improved Actuated Control Treatment

CORSIM output based on timing from:	Old T7F	New T7F
Control Delay	25.6	24.2
Total Delay	29.7	28.4
Stop Delay	22.0	20.5
Vehicle Trips	8116	8242

In summary, the objectives of this project are to improve the treatment of traffic-actuated control within deterministic models, to demonstrate the improvements using TRANSYT-7F release 8, and to validate the improvements using CORSIM. In support of these objectives, the following tasks will be carried out:

1. Identify specific shortcomings of the existing TRANSYT-7F actuated timing model:

   Related to this, generalized shortcomings of existing optimization models were discussed earlier in this chapter. Technical details regarding shortcomings of the existing TRANSYT-7F actuated timing model are provided in chapter 2.

2. Establish requirements for an improved model:

   To obtain an improved actuated timing model, it is necessary to keep in mind the minimum requirements for such a model. This prevents the possibility of wasted time in evaluating a new model that contains unacceptable weaknesses.

3. Survey candidate models for improvement of TRANSYT-7F performance:

   In the literature, there are a number of models that contain good ideas and methodologies, related to the subject matter at hand. It is necessary to survey these existing models, in order to learn whether existing methodologies may be helpful in developing a new model for improvement of TRANSYT-7F performance.

4. Evaluate candidate models per requirements:

   The minimum requirements, to be established by task #2, will be used to determine the existing models that are appropriate for further consideration.

5. Select one or more models for additional testing and development:

   Existing models that meet the minimum requirements for an improved model will be selected for additional analysis and consideration.

6. Develop models:

   Candidate actuated timing models must be developed for application in conjunction with TRANSYT-7F.

7. Formulate a test plan:

   Devise a comprehensive test plan to compile evidence of model effectiveness.

8. Test models per test plan:

   Testing will be performed on a variety of field conditions, in order to compile evidence on the effectiveness of the candidate models.

9. Recommend specific improvements:

   Conclusions and recommendations, regarding TRANSYT-7F actuated control modeling, will be formed based on the results from model testing.

Organization of Chapters

Table 1-2 illustrates the organization of chapters within this dissertation. It also shows which chapters will be used to address specific tasks and objectives listed earlier. The potential benefits of traffic signal timing optimization were briefly discussed in this chapter. Also, shortcomings of the existing TRANSYT-7F actuated timing sub-model were introduced, and the minimum requirements for an improved model were established. Thus, this chapter was used to address task 1. In support of tasks 2-5, chapter 2 provides background information on numerous models for traffic-actuated control. The third chapter is used to provide a frame of reference for each objective in this study. It describes technical aspects of the existing model including methodology, sample calculations, strengths and weaknesses, and preliminary test results (task 6). Chapter 4 outlines the testing plan, and contains the detailed testing results for the candidate models (tasks 7 & 8). The fifth and final chapter discusses implementation strategies, plus proposed additional testing and development of selected models (task 9).

Table 1-2: Study Objectives and Associated Chapters

Objective	Chapter
1. Identify shortcomings of the existing model	1 — Introduction
2. Establish requirements for an improved model	2 — Literature Review
3. Survey existing models	2 — Literature Review
4. Evaluate candidate models per requirements	2 — Literature Review
5. Select models for additional testing	2 — Literature Review
6. Develop models	3 — Model Development
7. Formulate test plan	4 — Model Testing
8. Test models per test plan	4 — Model Testing
9. Recommend specific improvements	5 — Conclusions and Recommendations

CHAPTER 2
LITERATURE REVIEW

This chapter summarizes the available literature on traffic-actuated control modeling, because such literature may be useful in determining the best strategies for improvement of TRANSYT-7F performance. Tasks 2 through 5 from chapter 1 (table 1-2) are addressed by this chapter, in which models from the literature will be summarized, evaluated, and possibly selected for further development. Because this discussion is primarily targeted at persons with experience in traffic operations, tutorial summaries on the basics of TRANSYT-7F and traffic-actuated control were moved from this chapter and are presented in the appendix.

Throughout the literature, it appears that models for traffic-actuated control can be classified into three major categories. This chapter contains three major sections for distinguishing between these model types. One of the basic requirements for an improved TRANSYT-7F model is that it must estimate average actuated phase times. Actuated phase time inaccuracy is the most significant perceived deficiency of the existing model. Because of this, models (in the literature) that are capable of estimating average actuated phase times may be especially useful. In addition, an improved model must be suitable for practical implementation within the TRANSYT-7F program.

Phase Time Estimation

Existing Model within TRANSYT-7F

Before investigating the candidate models for improvement of TRANSYT-7F performance, it is helpful to better understand the existing model. A candidate model must be considered superior to the existing model in order to warrant implementation, which is one reason that the existing model should be understood. Another reason is that one of the candidate models to be considered works similarly to the existing model. Finally, general understanding of the existing model facilitates general understanding of the candidate models and of actuated control.

The existing model for traffic-actuated control within TRANSYT-7F, developed at the University of California-Berkeley [Skabardonis, 1988], is primarily based on target degree of saturation. The model is designed to give actuated phases as little green time as possible, and gives all remaining green time to the major street through movement. How is it possible to give the actuated phases "as little green time as possible"?

It is necessary to give them a certain amount of green time that will result in near-saturation on that phase. This is consistent with typical operation, because the basic strategy of an actuated controller is to terminate the phase soon after the queue has been serviced. This results in a degree of saturation somewhat lower than 100%, although the exact degree of saturation is difficult to predict. The default value currently applied by the TRANSYT-7F program is 85%. Figure 2-1 illustrates the minor movements at 85% saturation.

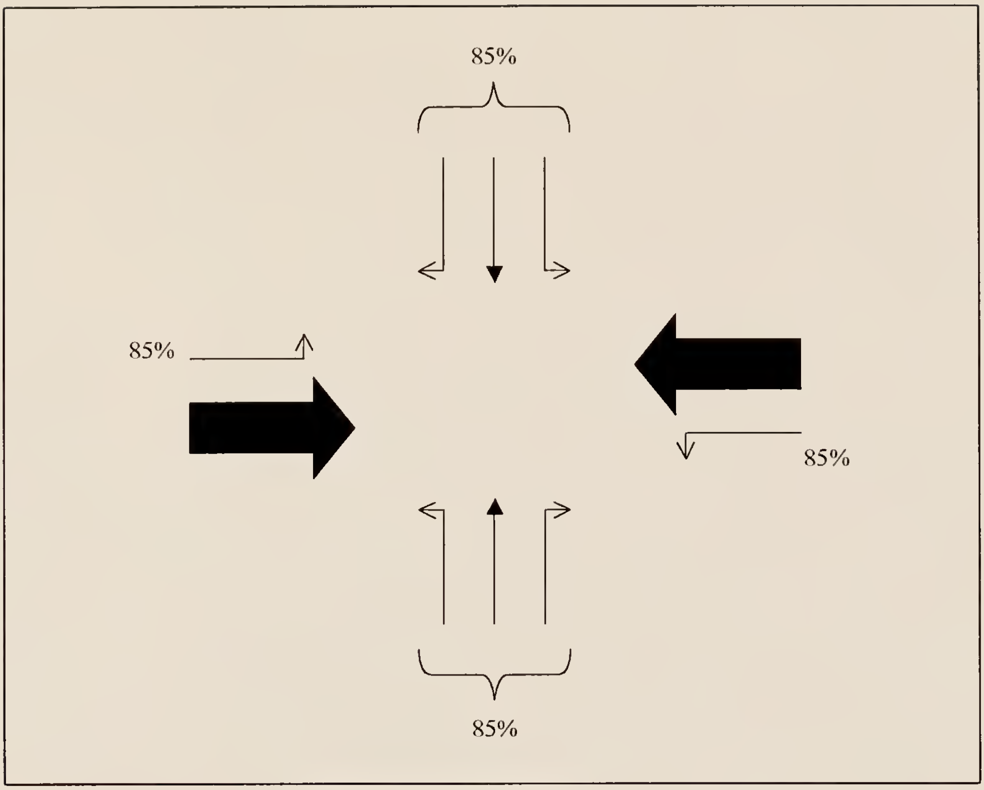


Figure 2-1: Existing Model Strategy — Actuated Phases at 85% Saturation

Degree of saturation concept

At this time it is necessary to define "degree of saturation." Mathematically, degree of saturation equals volume divided by capacity; or, $X = v/c$. Conceptually, if there is only enough green time (on average) to service the initial queue (on average), then that movement operates at 100% degree of saturation. In this case of 100% saturation, the expected volume (during each cycle) will receive just enough green time (during each cycle) to be served, and the capacity of the movement is equal to the expected volume. If there were not enough green time to service the initial queue, degree of saturation would exceed 100%. If there were enough green time such that the initial

queue could be served, plus some additional vehicle arrivals thereafter, degree of saturation would be less than 100%.

Suppose that the left-turning vehicles shown in figure 2-2 have just been given the green arrow to indicate the beginning of a protected left-turn phase. On the westbound approach, the initial queue contains six vehicles, whereas on the eastbound approach, the initial queue contains two vehicles. What if the eastbound approach were to receive enough green time such that the initial two vehicles, plus the two straggling vehicles behind them, were all served? Degree of saturation for that movement would be less than 100%. What if the westbound approach were to receive enough green time such that only four out of the six vehicles in the initial queue were served? The "residual queue" would be two vehicles, and the movement would be "oversaturated" with degree of saturation greater than 100%. In order to know the exact value of degree of saturation in any of these cases, it would be necessary to perform additional calculations regarding the value of capacity. This is illustrated later.

Volume (v)
Capacity (c)
Degree of Saturation ($X = v/c$)

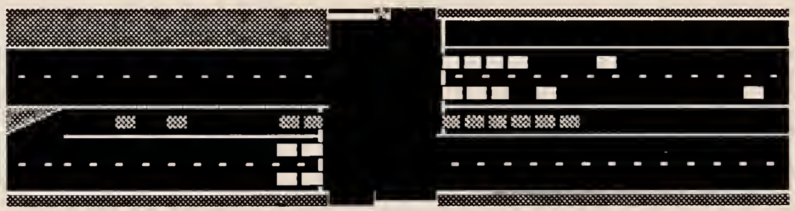


Figure 2-2: Degree of Saturation Concept Example Using Left-Turn Movement

Returning to the discussion about the existing TRANSYT-7F model, it was stated that the model is designed to give the actuated phases as little green time as possible, and gives all remaining green time to the major street through movement. Further, in order to give the actuated phases as little green time as possible, the model gives them enough green time such that their movements achieve 85% saturation. Why 85%?

In the real world, arrival rates and queues vary from cycle to cycle. Although the analyst typically knows the expected average arrival rate and average queue length, these values should be exceeded about half of the time. It would be highly undesirable to exceed 100% saturation during any given cycle, because when that occurs, vehicle delay tends to rise exponentially. In the existing model, target degree of saturation should be low enough such that 100% is rarely exceeded.

On the other hand, when target degree of saturation decreases for the minor movements, their allotted amount of green time increases. This conflicts with the desired objective of giving as much green time as possible to the major street through movement. Any unnecessary, wasted green time on the minor movements would be better spent on achieving arterial progression on the major street through movement. In the existing model, target degree saturation should be high enough such that green time is not wasted on the minor movements.

To summarize, in choosing target degree of saturation for an actuated phase, it is desirable to have it as high as possible, provided that 100% saturation is rarely exceeded. Skabardonis [1988] recommends target degree of saturation in the range of 85-90%. The value of 85% has been chosen as the default value in TRANSYT-7F, although the user is

allowed to specify any value between 50 and 100%, applicable to specific phases, or applicable to the entire network.

Sample calculation

To understand its relative effectiveness, it is helpful to walk through a sample calculation with the existing model. Figure 2-3 illustrates a set of hypothetical conditions at a signalized intersection.

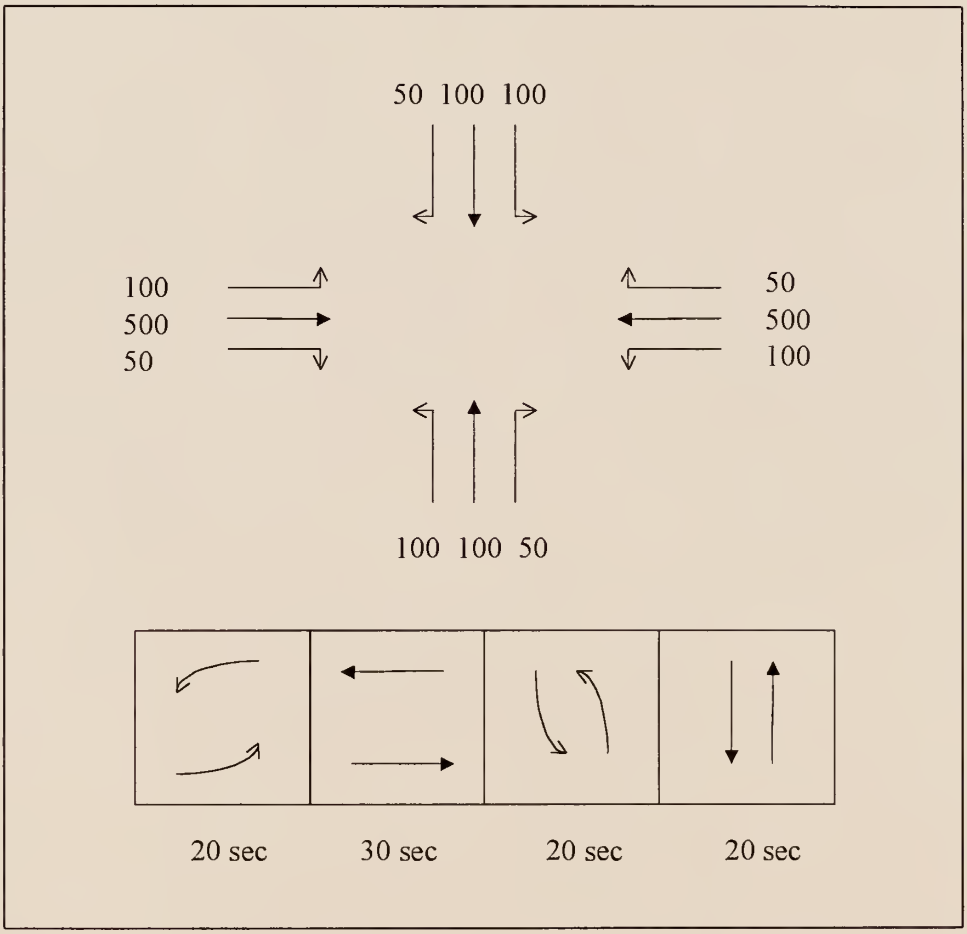


Figure 2-3: Existing Model Sample Calculation — Intersection Conditions

1. Calculate Capacity

Consider the north-south left-turn phase in the figure 2-3 timing plan. The largest individual movement volume on this phase is 100 vehicles per hour. In order to achieve 85% saturation, the phase must receive enough green time such that it could actually serve 118 vehicles per hour. Thus, movement capacity is equal to 118 vehicles per hour.

Volume (v) = 100 vehicles / hour

$$X = \frac{v}{c} = 0.85$$

$$c = \frac{v}{0.85} = \frac{100}{0.85} = 118 \text{ vehicles / hour}$$

Figure 2-4: Existing Model Sample Calculation Step #1

2. Calculate Required Percentage of Green Time per Hour

In step 1 it was shown that movement capacity must equal 118 vehicles per hour in order to achieve 85% saturation. Next, it is necessary to determine the amount of green time per hour needed to serve 118 vehicles on that movement. Suppose that the movement is capable of serving 1710 vehicles per hour if it could have an hour's worth of green time. If an hour's worth of green time results in 1710 vehicles served, how much green time results in 118 vehicles served? By simple math, if the movement received green time during 6.9% of the hour, then 118 vehicles would be served.

Capacity (c) = 118 vehicles / hour
Saturation Flow (s) = 1710 vehicles / hour green

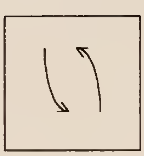
$$\frac{\text{Required Time}}{\text{Hour}} = \frac{\text{Capacity}}{\text{Saturation Flow}}$$

$$\frac{c}{s} = \frac{118}{1710} = 6.9\% \text{ of the hour is required}$$

Figure 2-5: Existing Model Sample Calculation Step #2

3. Calculate Phase Time

In step 2 it was shown that the movement must receive green time during 6.9% of the hour. This means that it must also receive 6.9% of the green time available during each signal cycle. The calculations illustrated show that, after adding on two seconds of



6.9% of the hour is required
Cycle Length (C) = 90 seconds
Total Lost Time (l) = 8 seconds

$$6.9\% = \frac{\text{green time}}{90 - 8}$$

$$\text{green time} = 82 \bullet 0.069 = 5.6 \text{ seconds}$$

$$\text{phase time} = 5.6 + 2 \approx 8 \text{ seconds}$$

Figure 2-6: Existing Model Sample Calculation Step #3

yellow and all red clearance time, the left-turn phase time is 8 seconds according to the existing model.

4. Donate Unused Green Time to the Non-Actuated Phase

In step 3 it was shown that the existing model estimates the average length of the first phase as 8 seconds. The same computations from steps 1-3 are performed to estimate the average length of each actuated phase. Subsequently, all remaining green time in the cycle is allocated to the non-actuated phase, as illustrated in figure 2-7.

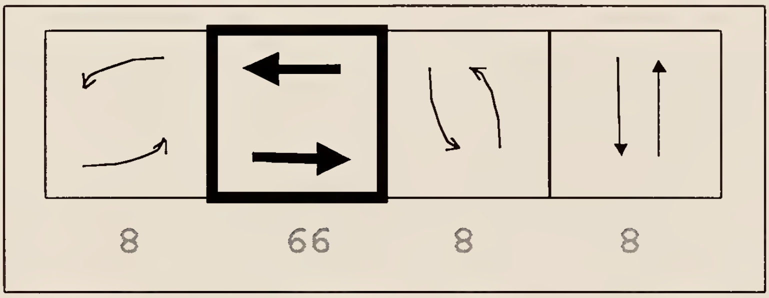


Figure 2-7: Existing Model Sample Calculation Step #4

The TRANSYT-7F output in table 2-1 presents the existing model sample calculation results. Along the rows labeled “Intvl Length” and “Splits,” the signal timing table shows that 63 seconds have been allocated to the coordinated, non-actuated phase. Along the row labeled “Phase Start,” the signal timing table shows that phase number 4 is the non-actuated phase (NAP). Along the row labeled “Links Moving,” the signal timing table lists the link numbers moving on each phase. For example, through and right-turn links (105, 107, 111, 112) are moving during the non-actuated phase. Along the column labeled “Deg Sat,” the measures of effectiveness table shows that simulated degree of

saturation was approximately 85% for the critical actuated movements (102, 104, 101, 103, 106, 108).

Table 2-1: TRANSYT-7F Program Output for Sample Calculation

```
-----------------------------------------------
 Intersection     1   Actuated - Splits Estimated
 -----------------------------------------------

 Interval Number  :      1    2    3    4    5    6    7    8

 Intvl Length(sec):    6.0  2.0 64.0  2.0  6.0  2.0  6.0  2.0
 Intvl Length  (%):      7    2   71    2    7    2    7    2

 Pin Settings  (%):  100/0    7    9   80   82   89   91   98

 Phase Start (No.):      1 AC T    2 NA P    3 AC T    4 AC T

 Interval Type    :      V    Y    V    Y    V    Y    V    Y

 Splits       (sec):     8        66         8         8
 Splits         (%):     9        73         9         9
 Links Moving     :    106       105       102       103
                       108       107       104       101
                                 111                 109
                                 112                 110

 Offset =    0.0 sec    0 %.

<PERFORMANCE WITH INITIAL SETTINGS>

 Movement/  Deg/  Total   Travel Time      Delay        Total    Max Back   Fuel
 Node Nos.  Sat  Travel   Total  Avg.   Total  Avg/LOS  Stops    of Queue   Cons.
             %    v-mi    v-hr  sec/v   v-hr    sec/v   No. (%)  Est.Cap.    gal
 ---------------------------------------------------------------------------------
   101     : 83   50.02   4.20 151.3   2.19   78.8E   140(140)   2    106    4.2
   102     : 88   50.02   4.52 162.9   2.51   90.4F   151(152)   2    106    4.5
   103     : 83   50.02   4.20 151.3   2.19   78.8E   140(140)   2    106    4.2
   104     : 88   50.02   4.52 162.9   2.51   90.4F   151(152)   2    106    4.5
   105     : 39  250.10  10.91  78.5   0.85    6.1A   190( 38)   4    106   12.6
   106     : 88   50.02   4.52 162.9   2.51   90.4F   151(152)   2    106    4.5
   107     : 39  250.10  10.91  78.5   0.85    6.1A   190( 38)   4    106   12.6
   108     : 88   50.02   4.52 162.9   2.51   90.4F   151(152)   2    106    4.5
   109     : 49   25.01   1.62 116.7   0.61   44.3D    50(101)   1    106    1.8
   110     : 49   25.01   1.62 116.7   0.61   44.3D    50(101)   1    106    1.8
   111     :  5   25.01   1.06  76.4   0.06    4.0A    14( 29)   0    106    1.2
   112     :  5   25.01   1.06  76.4   0.06    4.0A    14( 29)   0    106    1.2

 NODE     1: 88  900.37  53.69        17.46   34.9C  1392( 77)               57.4
 ---------------------------------------------------------------------------------
```

5. Subsequent Simulation or Optimization

Once the existing model has arrived at its estimates for phase times, the resulting timing plan is then used as a starting point in simulation or optimization. If optimization has not been requested whatsoever, then the phase time estimates will not be modified during simulation. If thorough optimization has been requested, then the phase time estimates are simply used as a starting point in a hill climb search for a better signal timing plan. If optimization of offsets-only has been requested, then TRANSYT-7F will not modify the phase time estimates as it searches for better offsets.

If the initial timing plan is close to the global optimum solution, then the hill climb procedure has a higher probability of locating that global optimum. Also, when there is more green time available for the coordinated through movement, then there is a higher probability of achieving progression. Thus, when the actuated model estimates phase times and allocates extra green time to the coordinated movement, TRANSYT-7F gets a better starting point for optimization of a congested artery.

Accurate phase time estimation can be critical. Inaccurate phase time estimation will result in the wrong amount of green time being allocated to the coordinated movement. Thus, the output measures of effectiveness and/or optimal offset timing produced by TRANSYT-7F may be overly optimistic or pessimistic. It depends on whether the amount of green time allocated to the coordinated movement, and the resulting available green band throughout the arterial, is too large or too little.

Specific shortcomings

Although effective at eliminating wasted green time, as an actuated controller would, the existing model for traffic-actuated control within TRANSYT-7F is

oversimplified. Predicted phase times are simply a function of the number of actuated phases and the target degree of saturation. The target degree of saturation strategy is inherently inaccurate because it is not responsive to numerous factors that affect actuated phase lengths. As stated in chapter 1, actuated phase times and measures of effectiveness reported by TRANSYT-7F are potentially less accurate, because the existing model is not sensitive to several key factors. In addition, since actuated phase times are often estimated prior to optimization, any phase time inaccuracies could lead to inferior optimization results. What are these key factors, and how do they affect actuated phase times?

Gap setting: If a phase is actuated, then its associated lanes contain detectors that search for a gap in the traffic stream. The gap setting indicates the size of the gap being searched for, measured in units of seconds. Larger gap settings produce larger actuated phase times. This is because large gaps in traffic occur less frequently than small gaps, and phases will continue to last longer if the desired gap is not detected. The existing model for actuated control within TRANSYT-7F gives the phase enough green time to achieve 85% saturation, as stated previously. Theoretically, for a given phase, there exists a gap setting that will produce a certain green time resulting in 85% saturation on average. Using this gap setting, the existing model would produce accurate phase time estimates. For example, suppose that someone in the field carefully observes the north and southbound left-turn phase from figure 2-3 for several cycles of operation. Suppose they conclude that the average phase time was indeed 8 seconds during that period of observation, just as predicted by the existing TRANSYT-7F model in table 2-1. For these conditions, the existing model is accurate.

But what if the associated gap setting were to be increased, and other factors and variables held constant? In the field, average phase time would be expected to increase along with the gap setting, thus reducing the average degree of saturation to somewhere below 85%. Because the analysts are unaware of the degrees of saturation associated with various gap settings, they are unable to estimate a target X other than 85%, which causes inaccurate phase time estimates. This is how a change in the gap setting can render the existing model less effective.

Detector configuration: As stated above, if a phase is actuated, then its associated lanes contain detectors that search for a gap in the traffic stream. The configuration includes the type (presence vs. passage), length, and lane location of detectors. Longer presence detectors lead to larger actuated phase times. This is because longer detectors have a better chance of detecting vehicles prior to a controller's irrevocable decision to terminate the phase.

For a given phase, a given presence detector length exists that, in conjunction with a given gap setting, will produce a certain green time resulting in 85% saturation on average. If this hypothetical detector length is exceeded and all other factors and variables are held constant, the resulting degree of saturation would fall somewhere below 85%. Thus, a change in detector length can also render the existing model less effective.

Force-off: As previously stated, if a phase is actuated, then its associated lanes contain detectors that search for a gap in the traffic stream. What if traffic happens to be heavy enough such that the desired gap is never detected? It is necessary to impose a maximum green time setting on the actuated phase. The phase must not be allowed to

continue for an unreasonable amount of time. When signals are coordinated, the force-off setting is used to terminate actuated phases that do not gap-out.

For a given phase, a range of force-offs exists that will allow green times resulting in 85% saturation on average. In other words, if the force-off setting is sufficiently high, the phase is capable of termination via gap-out, and 85% saturation may occur. But what if this force-off setting was decreased, and other factors and variables held constant? In the field, if the force-off occurred early enough in the cycle, the phase would be expected to terminate via max-out before locating the specified gap, and thus increase average degree of saturation to somewhere above 85%. Thus a change in force-off can also render the existing model less effective.

Physical factors: The previous subsections discuss the way in which individual actuated control parameters can affect average phase lengths. There are also some additional factors that may affect average phase lengths, according to the literature presented later in this chapter. These include vehicle length, number of lanes, and approach speed. The existing model for traffic-actuated control within TRANSYT-7F does not take these physical factors into consideration.

Operational characteristic effects: In the existing model, what if it were possible to make target degree of saturation sensitive to all of the control parameters and physical factors mentioned in the previous subsections? Would this fix the model? It would be a step in the right direction; however, there would still be no clear-cut way for dealing with operational characteristic effects on actuated phase times. The effects of early returns to green, overlap phasing, stochastic behavior, progression, permitted left-turns, spillback, and optimization, are based on performance at neighboring signals, and

are thus difficult to quantify without simulation. Chapters 3 and 4 illustrate the technical aspects of operational characteristic effects with more technical detail.

Trying to predict degree of saturation: In the field, actuated phase times result in a certain degree of saturation. If the user were able to predict degree of saturation based on available input data, it would be possible to effectively use the existing model by specifying the appropriate target degree of saturation for each phase. However, degree of saturation is too difficult to predict. Non-uniform arrival rates, link length, lane channelization, control parameters, and detector layout, for numerous phases and intersections, interact as significant variables at many levels, and thus conspire to produce complex and unpredictable results throughout the network. This is why programs such as CORSIM (to be described later) and TRANSYT-7F are needed. This is also why an alternative method for estimating actuated phase times, besides trying to forecast the degree of saturation, is desirable.

Pitfalls of inaccurate phase time estimation: A simple example is helpful for illustrating the potential pitfalls of inaccurate phase time estimation. The phasing diagram in figure 2-8 illustrates a hypothetical signal timing plan. This signal timing plan has leading left-turn phases with no overlaps. Average phase times are listed below the box that represents each phase. Under pre-timed control, phase times should be approximately equal for each phase in the event that traffic volume demands and saturation flow rates are approximately equal for each movement. In the four-phase situation illustrated by figure 2-8, phase times should be 25 seconds for each phase if the cycle length is 100 seconds. Although left-turns typically have lower volumes and saturation flow rates than through movements, suppose for this simple example that they

have equal volumes and saturation flow rates, and that phase times should be equal in order to optimize the situation.

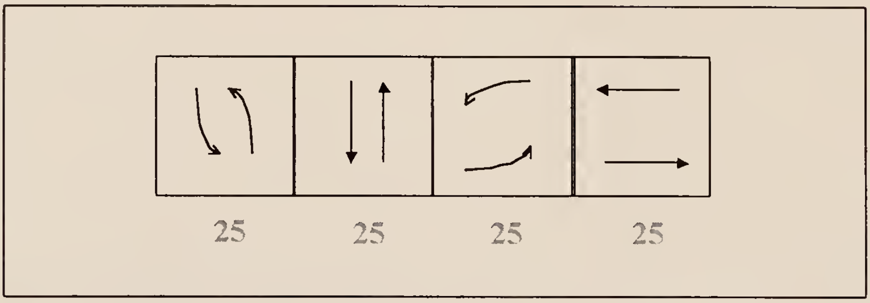


Figure 2-8: Sample Phasing Diagram with Average Phase Times

Now suppose that control type at this signal is to be converted to coordinated-actuated, with the east-west (left-to-right) through movement to be served by a coordinated phase. This means that the coordinated phase will have priority, and that the minor movement actuated phases are to be terminated as soon as possible after their initial queues have been served, such that the coordinated phase will receive extra green time. What will the new signal timing plan look like? Presumably something like the updated timing plan illustrated in figure 2-9. In this timing plan, actuated phase times are now much lower than their original pre-timed counterparts, and the coordinated phase benefits by receiving the extra green time. This type of control is typically preferable for achieving progression along the major street, provided that performance on the actuated phases does not deteriorate to unacceptable levels.

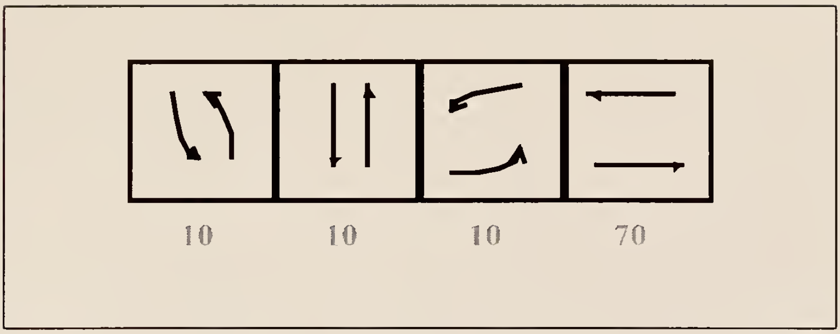


Figure 2-9: Sample Timing Plan Computed by the Existing Model

As mentioned earlier, accuracy of the existing actuated phase time model within TRANSYT-7F leaves much room for improvement. Figure 2-10 illustrates a hypothetical outcome of this situation. Actuated phase times that materialize in the field are actually 14 seconds apiece instead of 10, and the coordinated phase time occurring in the field is actually 58 seconds instead of 70. In the context of capacity analysis, the result would be an inappropriately pessimistic analysis, in terms of vehicle delay and level of service, of the actuated phases. Perhaps more dangerous, such results would cause overly optimistic analysis of the coordinated phase serving the major street movements. In the context of signal timing optimization, inaccurate phase time estimates can result in compromised estimates of the available green band, or green window, available for optimization of offsets or phasing to achieve progression along the major street.

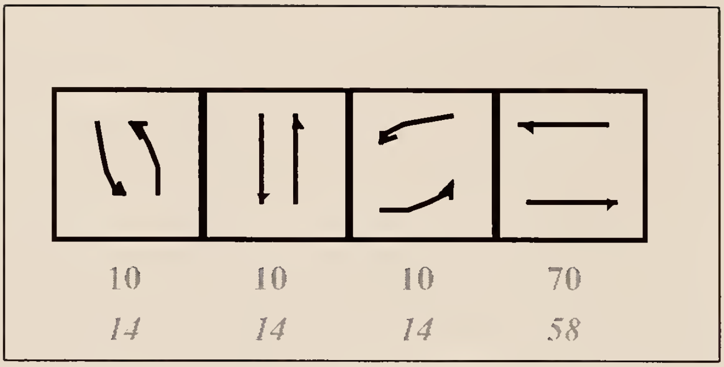


Figure 2-10: Example of Model-Predicted vs. Actual Actuated Phase Times

NCHRP Model

The model for estimating actuated phase times in appendix II of the 1997 Highway Capacity Manual is an example of a method that provides an alternative strategy to forecasting the degree of saturation. This model was developed as part of National Cooperative Highway Research Program (NCHRP) Project 3-48 [Courage et al., 1996]. An upcoming description of its methodology will be followed by a sample calculation using the same data from the existing model sample calculation (figure 2-3).

Methodology

In general, actuated phase lengths are estimated using the structure illustrated in figure 2-11. The NCHRP model specifies that the length of an actuated phase can be estimated by summing the queue service time (g_s), the green extension time (g_e), and the intergreen (yellow plus all red) time. Note that the green extension time should not be confused with the extension of effective green (EEG) parameter, which is applied by numerous deterministic traffic models. Although figure 2-11 does not illustrate start-up lost time, this parameter affects the phase time as well. According to the HCM terminology, queue service time begins where start-up lost time ends. However, in figure 2-11 and in other parts of this paper, queue service time includes start-up lost time.

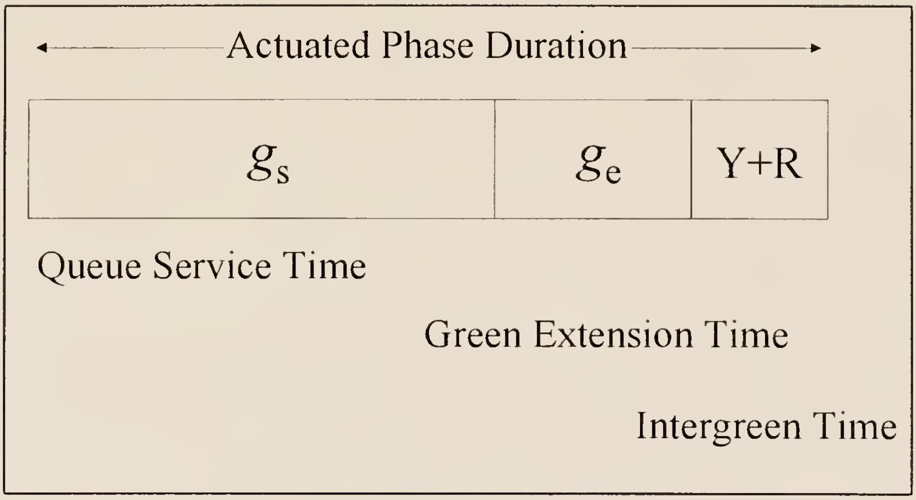


Figure 2-11: Individual Phase Length Structure within the NCHRP Model

Initial timing plan: The NCHRP model requires an initial timing plan in order to perform calculations. This initial timing plan makes it possible to determine the effective red time, and thus the expected initial queue, that is experienced by each actuated phase. Subsequently, the model is able to appropriately adjust the phase times in response to the differing characteristics of each phase, such as traffic volume demand and the actuated control parameters.

It is important to note that differing initial timing plans should not prevent the model from computing the same final solution each time. Only the input parameters affecting the model should impact the final solution; however, an initial timing plan that is close to the final solution results in fewer model iterations and faster running times for the computer program.

Queue service time: The queue service time for each phase is calculated using the queue accumulation polygon (QAP). The QAP concept is robust, reliable, and well-documented in the Highway Capacity Manual [Transportation Research Board, 1997]

and in other parts of the literature. The shape of the polygon is affected by several relevant parameters affecting vehicle queuing and delay, including traffic volume arrivals, saturation flow rate, effective red time, and effective green time. Figure 2-12 illustrates a sample QAP.

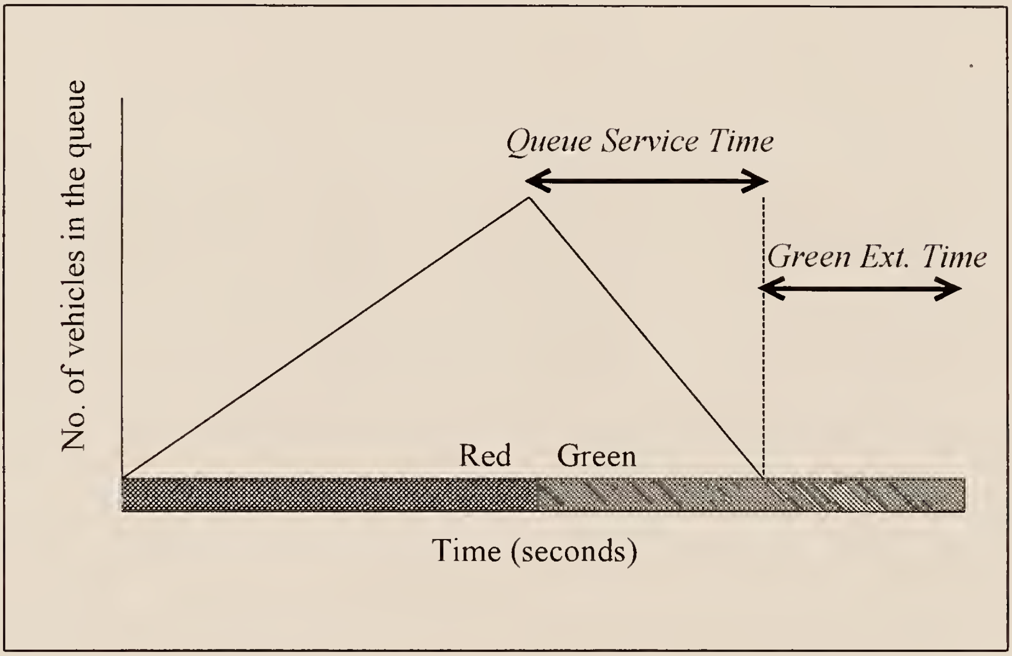


Figure 2-12: Sample Queue Accumulation Polygon

Figure 2-12 illustrates a polygon in the shape of a triangle, which implies that the rate of vehicle arrivals on red and the rate of vehicle departures on green are both constant, uniform rates. Unless traffic volume is very heavy, the slope of the left-hand side is not expected to be as steep as the slope of the right hand side, because when a queue is present, the rate of vehicle departures is very high (i.e., the saturation flow rate). The shape of the polygon is useful in determining queue service time among other things. The shape of the polygon can become complex in response to field conditions causing

non-uniform arrivals or departures, allowing queue service time to be computed accurately under complex conditions.

The HCM discussion lists a formula for computing queue service time. This formula (2-1) is an algebraic interpretation of the queue accumulation polygon, targeted at computing queue service time.

$$g_s = f_q \frac{q_r r}{(s - q_s)} \tag{2-1}$$

where,

q_r , q_g = red arrival rate (veh/s) and green arrival rate, veh/s, respectively,
r = effective red time, s,
s = saturation flow rate, veh/s, and
f_q = queue calibration factor

$$f_q = 1.08 - 0.1\left(\frac{\text{actual green}}{\text{maximum green}}\right) \tag{2-2}$$

Green extension time: The queue accumulation polygon is not as useful in determining the green extension time. Figure 2-12 illustrates that green extension time takes place beyond the polygon boundaries, and cannot be directly computed from the shape of the polygon. The HCM appendix states that green extension time for each phase can be calculated using equations developed by Akcelik [1993, 1994].

$$g_e = \frac{e^{\lambda(e_0 + t_0 - \Delta)}}{\varphi q} - \frac{1}{\lambda} \tag{2-3}$$

where

e_0 = unit extension time setting
t_0 = time during which the detector is occupied by a passing vehicle

$$t_0 = \frac{(L_d + L_v)}{v} \tag{2-4}$$

where

L_v = vehicle length, assumed to be 5.5 m
L_d = detector length, DL, m,
v = vehicle approach speed, SP km/h
Δ = minimum arrival (intra-bunch) headway, s,
φ = proportion of free (unbunched) vehicles, and
λ = a parameter calculated as:

$$\lambda = \frac{\varphi q}{1 - \Delta q} \tag{2-5}$$

where q is the total arrival flow, veh/s for all lane groups that actuate the phase under consideration. Akcelik [1999] has also developed formulas for computing actuated phase degrees of saturation. These formulas that compute degree of saturation will be presented later on in this chapter, but should not be confused with Akcelik's green extension time formulas above, which are an integral part of the NCHRP model.

Actuated phase lengths: The phase length is computed as the sum of the queue service time, the green extension time, and the yellow plus all red time. Yellow and all-red times are, of course, given parameter values that do not require calculation. Subsequently, the estimated average phase length is subject to the constraints of the associated minimum green, maximum green, and force-off settings. If the estimated phase length is lower than the minimum green time, it is set equal to the minimum green time. If the estimated phase length is higher than the maximum green value created by the force-off, it is set equal to that maximum green time.

Overall timing plan: The new set of estimated phase lengths associated with any given intersection may not sum to the background cycle length. This is not an issue under isolated intersection conditions, where the cycle length is simply recalculated as

the sum of the individual phase lengths. Under coordinated conditions, adjustments are necessary to conform the new solution to the background cycle length. If an estimated phase time indicates phase termination prior to reaching the force-off setting, then the unused green time must be reassigned within the cycle.

Assignment of unused green time is performed in accordance with the iterative computational structure described in the HCM guidelines. If an estimated phase length is lower than its original phase length from the initial timing plan, its new value is set equal to the midpoint between those two values. The other half of the slack time is then donated to the coordinated phase. For example, suppose the initial timing plan contains a 15 second actuated phase and a 45 second non-actuated phase. If the actuated phase is estimated as 9 seconds by the NCHRP model, then donating half of the slack time (instead of all of it) to the coordinated phase results in a 12-second actuated phase time and a 48-second non-actuated phase time for the second iteration. Donating half of the slack time to the coordinated phase, instead of donating all of the slack time, may increase the probability of convergence within the iterative model structure.

Iteration and convergence: After obtaining the overall timing plan, the new design must be tested for convergence. This is necessary because when the overall timing plan changes, it causes flow patterns to change throughout the analysis. Changing flow patterns result in changing phase lengths, and vice versa. This iterative process can be terminated when the timing plan on any iteration is observed to be identical to the one from any previous iteration. For example, if the timing plan in iteration 5 were identical to the one from iteration 3 or 4:

Table 2-2: NCHRP Model Intermediate Outputs

Iteration 1

Critlink	Start	MnG	Fq	Gs	Ge	Y+R	Split	Max	Min
106	0	4	.98	6.38	4.58	2	16	20	6
0	20	4	1	0	0	2	42	72	30
102	50	4	.98	6.38	4.58	2	16	20	6
103	70	4	.98	6.15	4.58	2	16	20	6

Sample calculation

A sample calculation is presented here to demonstrate an application of the NCHRP model. The initial timing plan (20 30 20 20) in iteration 1 reflects the same sample calculation conditions from figure 2-3, which was used earlier to demonstrate the existing model within TRANSYT-7F.

Individual phase times: Intermediate outputs provide details on how individual phase lengths are estimated. Consider the intermediate output from iteration 1:

Table 2-3: NCHRP Model Iterations of Overall Timing Plan

Iteration	Phase Times					
1	12	20	20	24	9	40
2	10	15	18	22	8	52
3	9	12	18	22	8	56
4	8	12	18	22	8	57

Convergence has been achieved

The start time within the cycle for north/south left-turn phase 3 (link 102) is time 50. Indeed, the search for a zero-length queue begins at the phase starting time. For phase 3, a zero-length queue must have been found after 6.38 seconds, because the queue

service time (g_s) is reported as 6.38 seconds. The green extension time (g_e) for phase 3 is reported as 4.58 seconds. This value was computed using Akcelik's formula. Yellow and all-red times for phase 3 are given for each phase as 2 and 0 seconds, respectively. Therefore, the actuated phase time for phase 3 in this iteration is computed as $g_s + g_e +$ yellow + all-red = 6.38 + 4.58 + 2 + 0 = 13 seconds (approximately). However, in the output file generated by the experimental version of TRANSYT-7F, “Split” is reported as 16 seconds for phase 3 in iteration 1. This is because half of the unused green time was donated to the non-actuated phase, in accordance with the NCHRP model procedures.

The estimated phase time of 16 seconds for phase 3 does not exceed the force-off time of 20 seconds and does not fall below the minimum phase time, reported in the output file as 6 seconds. Therefore, the estimated phase time does not violate either the minimum or the maximum constraints, and does not need to be readjusted prior to the next iteration.

Overall timing plan: After determining the each of the individual phase times in this manner, it is necessary to derive the overall timing plan to be used in the next iteration. The phase time for phase 3 was computed as 13 seconds. The NCHRP model calls for donation of half of the slack time. Because the left-turn phase lasted 20 seconds in the first iteration, it is reduced to 16 seconds in the second iteration, and 4 seconds are donated to the coordinated phase. Table 2-4 shows that in the first iteration, a total of 12 seconds is donated to the coordinated phase from the three actuated phases.

Thus, green times for the individual phases are reassigned in this manner until a brand-new timing plan has been established for use in the second iteration. At this point, the process would terminate if the brand-new timing plan were identical to any timing

plan from a previous iteration. In this case, the overall timing plan has changed and will be used in the second iteration. The new design is used to assist in determining the timing plan for the third iteration, and so on. In this case, convergence has been achieved after iteration number 7. The average phase times listed in the final line (14 48 14 14) are the ones that are expected to occur, given the input parameter values.

Table 2-4: NCHRP Model Timing Plan Iterations

Iteration	Phase Durations			
1	20	30	20	20
2	16	42	16	16
3	15	47	14	14
4	14	49	14	13
5	14	48	14	14
6	14	49	14	13
7	14	48	14	14

Existing model: Recall that the existing model for traffic-actuated control within TRANSYT-7F was also applied to analyze conditions from the sample calculation input file. The table 2-5 output reveals that phase times estimated by the existing model (8 66 8 8), are different than the ones produced by the NCHRP model (14 48 14 14).

Specific shortcomings

The NCHRP model is constrained by certain simplifying assumptions. The equations for green extension time listed earlier assume that detectors operate in the presence mode, and assume that detectors are installed at the stop line. Passage detectors, or detectors of any type that are not installed at the stop line, are not taken into account

by the formula for green extension time. In addition, computer execution time for the NCHRP model exceeds that of the existing model by many orders of magnitude, since numerous iterations are required.

Table 2-5: Existing Model Timing Plan for NCHRP Sample Problem

```
------------------------------------------------
Intersection     1   Actuated - Splits Estimated
------------------------------------------------

Interval Number  :       1    2    3    4    5    6    7    8

Intvl Length(sec):     6.0  2.0 64.0  2.0  6.0  2.0  6.0  2.0
Intvl Length  (%):       7    2   71    2    7    2    7    2

Pin Settings  (%):   100/0    7    9   80   82   89   91   98

Phase Start (No.):       1 AC T     2 NA P     3 AC T     4 AC T

Interval Type    :       V    Y    V    Y    V    Y    V    Y

Splits      (sec):       8        66         8         8
```

NCHRP model summary

The NCHRP model has the potential for improved accuracy over the existing model. It accounts for numerous elements that affect phase lengths, including actuated signal settings and detector layout. By varying these input conditions during testing, it is possible to show that the NCHRP model produces more accurate results than the existing model. Test results such as this are presented later on in this paper.

Iterative Target Degree of Saturation Model

Appendix II of the 1997 HCM states that there are two methods of determining the required green time given the length of the previous red. One of these methods is the NCHRP model, as described earlier. The second method describes the general procedure

for computing phase times as a function of the target degree of saturation. The existing model within TRANSYT-7F, also described earlier, employs this strategy. However, Akcelik [1999] has developed a unique model that combines these two strategies.

Earlier in this chapter it was stated that one of the specific shortcomings of the TRANSYT-7F existing model is its inability to respond appropriately to changes in traffic-actuated control parameters. Akcelik's model avoids this shortcoming to some extent by using special equations to compute the target degree of saturation. The equations specify that target degree of saturation is affected by the values of certain actuated control parameters, in addition to a couple of other relevant parameters.

Akcelik's model also involves elements of the NCHRP model. The NCHRP model procedure is iterative because of the circular dependency between individual actuated phase times. A change in phase length 'A' has an impact on the effective red time experienced by subsequent phase 'B'. This affects the average queue length at the beginning of phase 'B', which affects phase length 'B', possibly affecting the effective red time experienced by phase 'A', etc. Thus, the circular dependency between actuated phase times is recognized by dynamic changes in the queue service times.

In Akcelik's model, the circular dependency between actuated phase times is recognized by dynamic changes in the target degree of saturation. Two formulas are provided for calculating the target degree of saturation. The first formula is intended for use with initial calculations, when the effective red time is not known:

$$x_a = 1.5 y^{0.5} e_h^{-0.1} \tag{2-6}$$

subject to 0.40 <= x_a <= 0.95
where,

x_a = target degree of saturation for actuated signals
y = flow ratio (v/s)
e_h = effective headway (seconds)

The second formula is intended for use with subsequent iterations, when the effective red time is known:

$$x_a = 0.78 y^{0.5} e_h^{-0.1} r^{0.18} \tag{2-7}$$

subject to 0.40 <= x_a <= 0.95
where,
r = effective red time (seconds)

$$e_h = e_s + t_{ou} = e_s + \frac{3.6(L_v + L_p)}{v_{ac}} \tag{2-8}$$

where,
e_h = effective headway (seconds)
e_s = gap setting (seconds)
t_{ou} = detector occupancy time (seconds per vehicle)
L_v = average vehicle length (meters per vehicle)
L_p = effective detection zone length (meters)
v_{ac} = approach speed (km per hour)

Effective headway setting is a function of detector length and vehicle length, in addition to the gap setting. Thus, the target degree of saturation generated by this method is responsive to physical factors as well as signal settings. Since flow ratio and effective red time are also taken into consideration through the first two formulas, target degrees of saturation are appropriately sensitive to volume demand on the subject movement, in addition to phase lengths of the other movements. On the surface this appears to be an ideal candidate for upgrading the existing model without requiring substantial changes to the existing degree of saturation strategy. Instead of using a default value or a user input value for the target degree of saturation, this formula could be used to intelligently

predict degrees of saturation in response to numerous relevant parameters. However, earlier discussion on the existing model pointed out possible pitfalls of the target degree of saturation strategy. Specifically, there is no clear-cut way for revising the target degree of saturation strategy to account for added complexities such as protected-permitted left-turn phasing, or queue spillback.

In order to view a sample calculation according to the Akcelik's model, it is only necessary to review the earlier sample calculation for the existing model within TRANSYT-7F. Of course, the one exception to this is that target degrees of saturation for each phase would be calculated in advance via Akcelik's formulas, instead of using a default value or a user input value.

Modified NCHRP — Percentile Joint Poisson Probability Model

Husch [1996] describes a procedure for estimating actuated phase times that is similar to the NCHRP model. It utilizes the queue service time — green extension time strategy for calculating actuated phase times. However, Husch's model actually computes five hypothetical phase times, based on the 10^{th}, 30^{th}, 50^{th}, 70^{th}, and 90^{th} Poisson percentile versions of the average traffic volume. Subsequently, a volume-weighted average phase time is computed, based on the five hypothetical phase times. If any one of the five phase times violates the minimum or maximum phase time requirements, the average phase time can be adjusted accordingly. If the arrival rate on red for any one of the five hypothetical volumes is less than 0.69 vehicles, then the model assumes phase skipping (phase time = 0 seconds) for that volume scenario, based on a 0.5 Poisson probability of zero arrivals. Beyond this, the only clear difference between Husch's model and the NCHRP model lies in the calculation of green extension time.

Queue service time

Husch's queue service time formula is listed as equation (2-9) below.

$$T_q = T_s + \frac{A}{D-A} \times (Y_u + R + T_s) \tag{2-9}$$

where,

T_q = queue service time (seconds)
T_s = start-up lost time (seconds)
A = arrival rate (vehicles per hour)
D = saturation flow rate (vehicles per hour)
R = actual red time (seconds)
Y_u = unused yellow time (seconds)

This is essentially identical to the NCHRP model's formula for queue service time, presented earlier in this chapter as equation (2-1):

$$g_s = f_q \frac{q_r r}{(s - q_s)} \tag{2-1}$$

Although there are some apparent differences in the formulas, these differences are mostly cosmetic, and would not produce any difference within the overall results. The two formulas are essentially identical because the term (Y_u + R + T_s) is the same as effective red (r), arrival rate A is the same as arrival rate q, and departure rate D is the same as saturation flow rate s. Husch's formula contains start-up lost time (T_s), which is not listed in the NCHRP queue service time formula. However, NCHRP procedures stipulate that start-up lost time must be added to the queue service time and green extension time. The question of whether to define start-up lost time as a separate entity, or as part of the queue service time, is purely semantic and would not affect results.

Some differences are visible that could potentially introduce a bias into the results. The NCHRP formula contains two variables, q_g and q_r, to represent vehicle arrival rates on green and red, respectively. Husch's formula contains only one variable, A, to represent the arrival rate. In addition, Husch's formula does not implement the queue calibration factor, f_q, which accounts for stochastic behavior. Therefore, in situations where the arrival rate on green differs from the arrival rate on red, or in comparing the results to those from stochastic simulation, the NCHRP formula would presumably produce better correlation.

Green extension time

Husch's green extension time model is completely different than the NCHRP green extension time model (Akcelik's formula). Husch's model stipulates that once the queue has been serviced, phase termination via gap-out occurs when there is at least a 0.5 Poisson probability of zero arrivals. The Poisson calculations are performed for each second within the cycle, following the queue service time. For example, given a flow rate of 500 vehicles per hour, the Poisson probability of zero arrivals in the second that immediately follows queue service time is:

$$P(0) = e^{-\left(\frac{500}{3600}\right)} = 0.87$$

Given a gap setting of 3 seconds, the joint probability of three consecutive seconds of zero arrivals is:

$$P(0,0,0) = 0.87 \times 0.87 \times 0.87 = 0.66$$

Since the probability of three consecutive zero arrivals is greater than 0.5, an immediate gap is assumed, and the green extension time is equal to 3 seconds. Under a

different set of input conditions, if an immediate gap were not found, then the determination of whether a greater than 0.5 probability of gap-out occurs on subsequent seconds changes slightly.

Table 2-6 demonstrates the relationship between traffic volume and gap setting. This model computes an effective gap setting (GapEff) as a function of the actual gap setting and the detector layout. This allows phase times to be computed under various detector configurations. However, for standard presence detectors installed at the stop line, the effective gap setting would still be equal to the actual gap setting. In table 2-6, a cell value of zero indicates that the phase was not extended, and that the green extension time is equal to the gap setting.

Analysis

Although the volume-weighting approach may be useful in calculating an accurate average phase time, the individual phase times calculated for any given volume may be underestimated by Husch's model. First, by omitting the queue calibration factor, f_q, the queue service time model is unable to increase phase times in response to stochastic effects. Second, by assuming phase termination whenever there is a better than 50% chance of gap-out, the green extension time model is unable to increase phase times in response to vehicle extensions of the phase. In other words, if there were a 60 percent chance of gap-out after 3 seconds, a 20 percent chance of gap-out after 6 seconds, and a 20 percent chance of max-out after 9 seconds, this model would nevertheless compute a green extension time of 3 seconds. However, the appropriate green extension time in this scenario, considering each possible outcome, would actually be:

$$g_e = (0.6 \times 3) + (0.2 \times 6) + (0.2 \times 9) = 4.8 \text{ seconds}$$

Table 2-6: Husch's Green Extension Time Model Matrix

Volume (vph)	GapEff = 2	GapEff = 3	GapEff = 4	GapEff = 5	GapEff = 6
500	0	0	0	1	1
700	0	0	1	2	3
1000	0	1	2	4	6
1500	1	3	6	10	17
2000	2	5	12	23	49
2500	3	9	22	48	101

The mathematics of this model make it appear as though it would tend to underestimate actuated phase lengths, and subsequently overestimate green time available on the coordinated, non-actuated phase. Nevertheless, more detailed analysis and testing will be conducted on Husch's green extension time model later on in this study, in order to confirm these appearances.

Other aspects of Husch's overall phase time model need not be tested for the purposes of this study. The queue service time model is no different than that of the NCHRP, except that Husch's model does not incorporate the queue calibration factor and has fewer variables to describe the arrival rate in detail. The technique of performing calculations on five different Poisson percentile arrival rates seems only appropriate for external links, or at isolated intersections. The assumption of a Poisson distribution

would invalidate the results on internal actuated links with a nearby upstream signal, where vehicle arrivals are clearly non-random.

EVIPAS Program

Similar to TRANSYT-7F, the EVIPAS program was developed for the purpose of traffic signal timing optimization. The primary differences between the two programs are as follows. The primary feature of EVIPAS is that it was originally designed to optimize traffic-actuated signal settings such as the maximum green and the gap setting. The primary limitation of EVIPAS is that it is only capable of analyzing isolated intersections. The following description of the program appears in a paper titled "Design and Optimization Strategies for Traffic-Actuated Signal Timing Parameters" [Hale, 1995]:

> EVIPAS (Enhanced Value Iteration Program for Actuated Signals) is a program that has the ability to explicitly simulate traffic-actuated control. Its unique characteristic is the ability to optimize pre-timed or traffic-actuated signal settings by performing large quantities of iterative simulation runs. EVIPAS performs event-based simulations (the events being green extensions or green terminations) of an isolated intersection [Halati, 1992]. Since the simulations are event-based, they would tend to have much faster run times than individual CORSIM microscopic simulation runs for identical conditions. However, EVIPAS is only capable of modeling a single, isolated intersection. CORSIM and TRANSYT-7F have the ability to model arterials or networks with multiple intersections.
>
> EVIPAS performs the iterative simulation runs while attempting to minimize internally calculated vehicle delay. Univariate and gradient search techniques are used to find the optimal signal settings for an intersection with numerous input

field characteristics. The user can request to optimize any of the available traffic-actuated signal settings for a given controller. The user can also request to hold constant any of the signal settings in any phase and to optimize the others. The reduced run time for a single simulation run allows the optimization routine to perform hundreds or thousands of iterative simulation runs in a reasonable amount of time.

The EVIPAS program is unique to this study in the sense that it belongs within all three major categories of actuated control models in the literature (phase time estimation, delay estimation, and control parameter design/optimization). In the context of this literature review section on phase time estimation, EVIPAS appears to be an attractive candidate model for inclusion into, or evaluation of, TRANSYT-7F. It not only reports average phase times as outputs, but it has the ability to explicitly simulate actuated control with reasonable computer execution time. However, the current structure of EVIPAS, which assumes isolated operations, does not allow a background cycle length to be applied during the analysis. This existing limitation makes the program less useful in the context of evaluating or enhancing TRANSYT-7F network analysis of actuated control. Hopefully future versions of EVIPAS will have the added capability of optimizing traffic-actuated signal settings when a background cycle length is in effect.

CORSIM Program

The CORSIM program explicitly simulates traffic-actuated control. The same logic employed by actuated controller hardware in the field has been embedded within CORSIM. Therefore, CORSIM can be used as a tool in evaluating the candidate models for improvement of TRANSYT-7F performance. Unlike TRANSYT-7F, CORSIM (once

known as TRAF-NETSIM or NETSIM) was not developed for the purpose of signal timing optimization. It was meant to be an evaluation tool that could simulate a wide variety of traffic conditions.

Like TRANSYT-7F, the NETSIM component of CORSIM is a relatively old program that has earned a certain degree of recognition and acceptance within the transportation profession. The FRESIM component of CORSIM, used for simulating freeway links, is relatively newer. Since CORSIM was meant to be an evaluation tool, a fair amount of research has been done involving comparisons between its results and observed field data. One example of such comparisons involving field-measured traffic-actuated phase times is illustrated in figure 2-13 [Courage et al., 1996]. The realism of simulated traffic-actuated phase times is important in the context of this study.

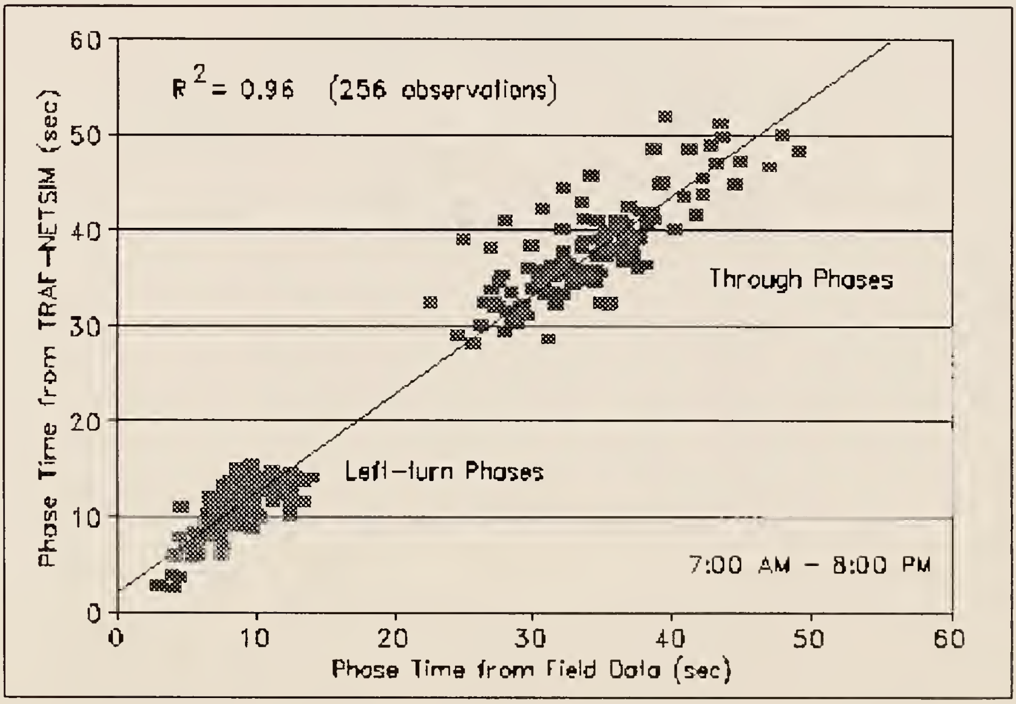


Figure 2-13: Traffic-Actuated Phase Times — CORSIM vs. Field Measured

The following description of the program appears in the CORSIM User's Manual [ITT Systems & Sciences, 1998]:

> CORSIM applies interval-based simulation to describe traffic operations. Each vehicle is a distinct object that is moved every second. Each variable control device (such as traffic signals) and each event are updated every second. In addition, each vehicle is identified by category (auto, carpool, truck, or bus) and by type. Up to 9 different types of vehicles (with different operating and performance characteristics) can be specified, thus defining the four categories of the vehicle fleet. Furthermore, a "driver behavioral characteristic" (passive or aggressive) is assigned to each vehicle. Its kinematic properties (speed and acceleration) as well as its status (queued or moving) are determined. Turn movements are assigned stochastically, asare free-flow speeds, queue discharge headways, and other behavioral attributes. As a result, each vehicle's behavior can be simulated in a manner reflecting real-world processes.
>
> Each time a vehicle is moved, its position (both lateral and longitudinal) on the link and its relationship to other vehicles nearby are recalculated, as are its speed, acceleration, and status. Actuated signal control and interaction between cars and buses are explicitly modeled.
>
> Vehicles are moved according to car-following logic, response to traffic control devices, and response to other demands. For example, buses must service passengers at bus stops (stations); therefore, their movements differ from those of private vehicles. Congestion can result in queues that extend throughout the length of a link and block the upstream intersection, thus impeding traffic flow.

In addition, pedestrian traffic can delay turning vehicles at intersections.

The above description of CORSIM highlights some of the fundamental differences between it and TRANSYT-7F. Unlike TRANSYT-7F, which is mesoscopic, CORSIM is a microscopic model in which "each vehicle is a distinct object" [ITT Systems & Sciences, 1998]. Moreover, CORSIM is a stochastic model: "Turn movements are assigned stochastically, as are free-flow speeds, queue discharge headways, and other behavioral attributes" [ITT Systems & Sciences, 1998]. The detailed nature of CORSIM also leads to significant input data requirements and relatively long execution times on the computer.

Since it recognizes actuated control parameters as inputs and produces phase times as outputs, CORSIM does meet the minimum requirements as a model that may improve TRANSYT-7F performance. The execution time of CORSIM prevents simultaneous processing with TRANSYT-7F. As realistic as CORSIM is, it should at least be used as a key evaluation tool in this study. It can be used to evaluate the candidate models for improvement of TRANSYT-7F performance.

The model testing experiments in chapter 4 demonstrate that variation of relevant input parameters within CORSIM tends to cause the appropriate changes in the actuated phase times. Variation of the same input parameters within TRANSYT-7F does not cause the appropriate changes in the actuated phase times, under the existing model. When the same input conditions are specified within TRANSYT-7F, comparisons between the two programs become possible.

Figure 1-1 from Chapter 1 illustrates mediocre correlation of actuated phase times calculated by CORSIM and the existing model within TRANSYT-7F. Similar

experiments later on in this study should indicate that better correlation is possible by implementing upgraded sub-models for actuated control within TRANSYT-7F. The appendix describes techniques for calibrating TRANSYT-7F and CORSIM to achieve better general agreement between the two programs; however, calibration efforts are not capable of significantly improving actuated phase time estimates from the oversimplified TRANSYT-7F existing model.

Phase Time Estimation Summary

To this point, the literature review has focused on models with the ability to estimate average phase times under traffic-actuated control. In the context of this study, these models are important to consider while attempting to develop an improved treatment of actuated control within TRANSYT-7F. The next two sections of literature review focus on the other classes of actuated control models, namely vehicle delay estimation and control parameter design and optimization. These two sections are useful for gaining perspective on the overall study; however, the upcoming models will not be immediately applicable for developing an improved model as described in the upcoming chapters.

Vehicle Delay Estimation

In the literature, models that calculate vehicle delay on an actuated phase have the tendency to predict better performance for that phase, compared to a non-actuated or pre-timed phase of equal duration. This is understandable because an actuated phase has a lower probability of temporary oversaturation, due to its ability to respond to variations in

traffic flow. These models in the literature that calculate vehicle delay on actuated phases will not, in the context of this study, be candidates for improvement of TRANSYT-7F performance. However, in order to better understand the candidate models for estimating phase times, it may be useful to understand how these delay calculators work. In fact, TRANSYT-7F currently implements one of these models for calculating vehicle delay on actuated phases. It currently implements the 1997 HCM delay equation [Transportation Research Board, 1997].

HCM Actuated Delay Model

The Highway Capacity Manual (HCM) was written for the purpose of evaluating the performance of various highway facilities. Similar to CORSIM, it was originally designed to be an evaluation tool, having no algorithms directly associated with design or optimization. Unlike the computer programs mentioned in this chapter, the HCM has no capability for simulating the flow of traffic within a simulated intersection or network. Rather, it uses various tables and equations to estimate the capacity of a given facility; although in some cases simulation programs may have been used to assist in developing certain tables and equations. It also reports the associated "level of service" along with other measures of effectiveness for the given facility, based on user inputs and intermediate computations. Table 2-7 [Transportation Research Board, 1997] is an example of one of these tables.

The HCM has been revised and published several times over the past few decades. At the time of this writing, the next update (HCM 2000) is scheduled to be published in October of the year 2000. The manual is a true benchmark within the transportation profession, and has earned a high degree of recognition. Although the

HCM is a manual or document, there are numerous software packages that incorporate portions of its procedures, including TRANSYT-7F. Delay equations are implemented within multiple HCM procedures, including the one for signalized intersections.

Table 2-7: HCM Level of Service Criteria for Signalized Intersections

Level of Service	Control Delay / Vehicle (sec)
A	0 - 10
B	10 - 20
C	20 - 35
D	35 - 55
E	55 - 80
F	80 +

The current (1997) delay model for signalized intersections begins with a large equation, having three "terms". These three terms contain numerous individual variables that are affected by many other tables, equations, and sub-models within the procedure. The first term of the delay equation is used to calculate uniform delay. Generally speaking, uniform delay can be calculated by measuring the area inside the queue accumulation polygon (QAP) illustrated earlier in figure 2-12. Conceptually, this is the vehicle delay that would occur if traffic flow were not stochastic, that is; identical queues and green times on each cycle, identical performance by each driver and vehicle, etc. Therefore, the first term of the delay equation is not responsive to actuated control, which is a stochastic process. Uniform delay is typically unchanged when changing between actuated and non-actuated phase definitions, provided other inputs are held constant.

The third term of the delay equation is used to calculate residual delay. This only becomes applicable when queues of unserved (by the previous green phase) vehicles are present at the beginning of the analysis. When conditions are temporarily oversaturated,

actuated phases are typically driven to their maximum green time on each cycle, and thus behave like non-actuated phases. Therefore, the third term of the delay equation is not responsive to actuated control. Residual delay is typically unchanged when changing between actuated and non-actuated phase definitions, provided that all other input conditions are held constant.

The second term of the delay equation is used to calculate incremental delay. The HCM states that incremental delay occurs "due to non-uniform arrivals and temporary cycle failures (random delay) as well as that caused by sustained periods of oversaturation (oversaturation delay)." Because traffic-actuated control is effective at reducing the probability of temporary cycle failures, the second term of the delay equation is indeed responsive to this. Incremental delay typically decreases when changing between actuated and non-actuated phase definitions, provided that all other input conditions are held constant. This delay term is listed as equation (2-10).

$$d_2 = 900T\left[(X-1)+\sqrt{(X-1)^2+\frac{8kIX}{cT}}\right] \tag{2-10}$$

where

T = duration of analysis period, hours;
k = incremental delay factors that is dependent on controller settings;
I = upstream filtering/metering adjustment factor;
c = lane group capacity, vph;
X = lane group v/c ratio, or degree of saturation.

TRANSYT-7F currently implements the 1997 HCM delay equation. However, performance of the traffic flow simulation model has an impact on the estimated delay, because the values of several variables within the equation are computed based on the simulation results. The values of the other variables are obtained directly from user

input. The end result is that TRANSYT-7F usually computes the same vehicle delay as other programs implementing the HCM procedures. Differences in estimated delay will sometimes be introduced due to the differences within other sub-models. For example, TRANSYT-7F by default applies an Australian model to compute permitted left-turn capacities, whereas the HCM signalized intersection procedures contain their own unique permitted left-turn movement model. All in all, TRANSYT-7F is considered up-to-date in the way that it calculates vehicle delay, thanks to the HCM delay equation.

It so happens that one of the variables that has a huge impact on the results of the delay equation is the average phase time. This underscores the importance of upgrading the actuated phase time calculation methodology that is currently implemented by TRANSYT-7F. Improvements to the accuracy of phase time calculation will automatically improve the accuracy of vehicle delay calculation, without even changing the way vehicle delay is calculated.

Volume-Weighted Average

In order to calculate vehicle delay at actuated signals, Husch [1996] recommends implementing the delay equation five times in a row based on five hypothetical traffic volumes, and then using a volume-weighted average to calculate the overall average delay. The hypothetical traffic volumes are estimated by assuming random Poisson arrivals that are often observed on the minor street, and then calculating the expected 10th, 30th, 50th, 70th, and 90th percentile volumes. This allows for some actuated control related adjustments, depending on whether some of these volume levels would clearly cause the phase length to violate the minimum or maximum phase times. The equation for calculating the volume-weighted average is listed below:

$$D = \frac{D10*V10 + D30*V30 + D50*V50 + D70*V70 + D90*V90}{V10 + V30 + V50 + V70 + V90} \qquad (2\text{-}11)$$

CORSIM and EVIPAS Programs

CORSIM computes and reports three types of vehicle delay. Instead of using an equation to estimate delay, it is computed directly from simulation. Total delay is determined by CORSIM as the difference between actual travel time and the travel time if constantly moving at the free flow speed. Queue delay is accumulated for each second when a vehicle has an acceleration of less than 2 feet per second squared, and a speed of less than 3 feet per second. Queue delay is accumulated every other second when a vehicle has an acceleration of less than 2 feet per second squared, and a speed between 3 feet per second and 9 feet per second. Stop delay is accumulated for each second when a vehicle has a speed of less than or equal to 3 feet per second.

Like CORSIM, EVIPAS computes delay directly from simulation. Exact measurement criteria similar to those listed above for CORSIM are unknown for EVIPAS. EVIPAS is therefore mentioned here so that all known models for calculating traffic-actuated vehicle delay are listed in this chapter, and to document the fact that EVIPAS is not known to use a separate analytical model for calculating delay, as does TRANSYT-7F.

Vehicle Delay Estimation Summary

This section of literature review has focused on models with the ability to estimate vehicle delay under traffic-actuated control. Improved vehicle delay estimation is not the direct objective of this study. It is important to note that an improved

methodology for calculation of actuated phase times should automatically result in better estimates of delay and other measures of effectiveness. The upcoming final section of literature review focuses on actuated control models for control parameter design and optimization. Again, this section is useful for gaining perspective on the overall study; however, the upcoming models will not be immediately applicable for developing an improved model as described in the upcoming chapters.

Control Parameter Design and Optimization

As mentioned earlier in this chapter, the requirements of an improved model include the ability to estimate average actuated phase lengths as outputs. Actuated phase lengths cannot be directly designed or optimized. They must instead be calculated as a function of numerous other parameters, some of which are indeed subject to design and optimization.

The models presented in this subsection were instead developed to design actuated control parameter values as outputs. The possibility exists that TRANSYT-7F could someday assist in directly designing actuated signal timing parameters. However, that is not the purpose of this study. These models presented in this subsection may provide a frame of reference. Although the overall objective of this study is to provide an improved methodology for calculation of actuated phase times, this methodology may result in a new facility for designing actuated control parameters within TRANSYT-7F.

Minimum Green Design

Some traffic-actuated signal settings, including the minimum green, cannot technically be optimized. The minimum green time should be set to a value that accommodates pedestrian traffic and driver expectations. Regardless of controller type, the value of the minimum green is generally at least 5 seconds for left-turn phases and 10 seconds for through phases, just to satisfy driver expectancy. The minimum green is increased and designed as a function of walking speed if pedestrians would require the extra time for crossing the intersection. These practices illustrate that design of the minimum green setting is governed by safety considerations.

Gap Setting Design

Engineering judgment is necessary in designing the gap setting (a.k.a. "unit extension"). If the driver of a queued vehicle is not paying attention when the queue begins moving, the detector may sense a gap and terminate the phase prematurely. A low-value gap setting can lead to premature phase terminations and excessive delays on an actuated phase unless the drivers are very attentive. On the other hand, a high-value gap setting can lead to wasted green time on the actuated phase, leading to excessive delays on the other phases. Gap settings can thus be improper if they are too low, or too high.

Several methods for design of the gap setting exist in the literature. Most methods are jurisdiction-specific recommendations of a certain value or a narrow range within x and y seconds. One method [Courage and Luh, 1989] recommends gap settings based on a statistical analysis that contains approximations. It assumes that a phase should be terminated when it may be concluded with 95 percent confidence that the

current flow rate has dropped below 80 percent of the saturation flow rate. EVIPAS was designed to optimize the gap setting, but this feature may be inappropriate from a practical standpoint, since the program contains no methodologies for modeling inattentive drivers.

Computer programs such as CORSIM, EVIPAS, and TRANSYT-7F typically do not model inattentive drivers or premature phase termination. The engineer should thus select a gap setting with the prospect of premature phase termination in mind, and then define that value in the programs, realizing that the models will not simulate inattentive drivers.

Maximum Green Design and Optimization

Numerous design and optimization methods for maximum green exist in the literature. They include a mixture of charts, programs, graphs, tables, and guidelines. Each one is a function of a different set of field variables.

The method for design of the maximum green setting described in the Methodology for Optimizing Signal Timing (MOST) manual [Courage and Wallace, 1991] can be applied through the SOAP program. The equations used by SOAP to calculate its maximum phase times involve computation of a traffic-actuated cycle length based on a user-specified target volume-to-capacity (v/c) ratio, and calculation of the phase times to be proportional to the critical flow ratios for each phase. This method is similar to the existing model for traffic-actuated control within TRANSYT-7F.

The method developed by Lin [1985] normally performs calculation of the maximum green setting as a function of the peak hour factor and the optimal pre-timed green setting. However, this method also incorporates compensation for the extra time

that is required for exclusive right-turn lanes. The formulas recommended by Skabardonis [1988] and modified by Fambro et al. [1992] are sensitive to the volume-to-capacity ratio, and recognize the impact of oversaturation on the design of the maximum green setting. Apart from their individual reference sources, the equations associated with these techniques for maximum green and gap setting design are compiled in a paper titled "Design and Optimization Strategies for Traffic-Actuated Signal Timing Parameters" [Hale, 1995].

Detector Configuration Design

Appropriate detector configuration design is a function of numerous characteristics. Presence detectors are typically installed near the stop line in order to detect the first vehicle in queue (see figure 2-14). Detector length specification is another engineering judgment call. Detector lengths that are too short increase the risk of vehicles going undetected, and increase the risk of premature phase termination. Detectors that are too long become expensive to purchase, install, and maintain. Also, if premature phase termination does not occur, longer detectors may increase vehicle delay. Multiple detectors within the same lane can affect traffic performance.

Evidence [Cribbins, 1975] [Lin, 1985] supports the intuitive notion that presence detectors that are too long lead to wasted green time and increased delay. In Florida, presence detectors are typically 30 feet long. Orcutt [1993] recommends 40-foot presence detectors, but adds that 60 foot detectors are preferable when approach speeds exceed 35 mph. Passage or pulse detectors are recommended if all traffic to be detected will flow freely between the detector and the intersection. Because they are much smaller than presence detectors, they are generally less expensive to install and maintain.

Detector length is not the primary design consideration as with presence detectors. Instead, detector setback is. As illustrated in figure 2-14, setback is the distance between the stop line and the detector.

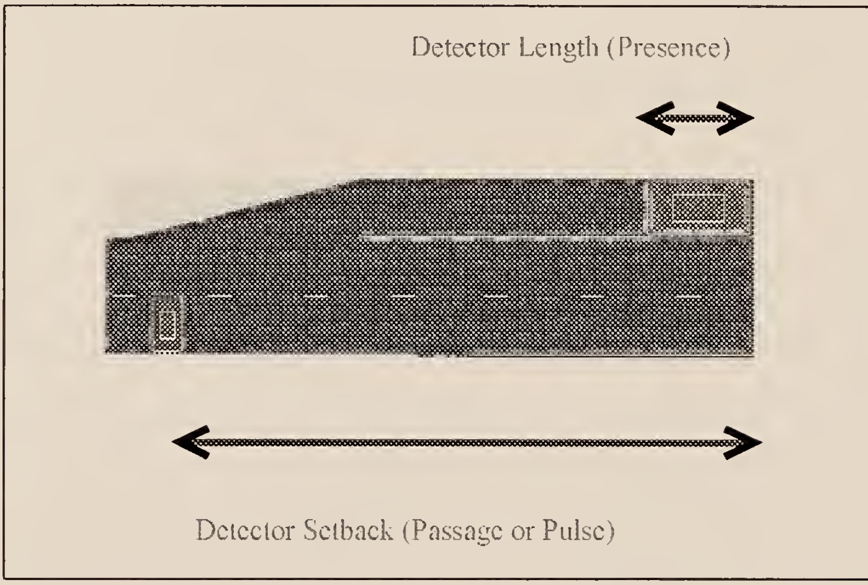


Figure 2-14: Sample Detector Layout or Configuration

Several sources [Kell and Fullerton, 1991] [Lin, 1985] [Tarnoff and Parsonson, 1981] recommend that setback should be designed based on safe stopping distance, which is a function of approach speed. The main purpose is to avoid the dilemma zone in which a vehicle can neither pass through the intersection nor stop before the stop line [Lin, 1994]. Bonneson and McCoy [1993] propose the alternative strategy of carrying the last clearing vehicle only through the indecision zone (rather than into the intersection) upon gap-out. Other sources [Bonneson and McCoy, 1993] [Fambro et al., 1992] describe the added benefit of using multiple advance loops in the detector design.

Coordination Setting Design

Certain traffic-actuated signals settings, such as the yield point, force-off, and permissive period settings, are applicable only in coordinated, actuated systems. However, some models recommend optimal values using the terminology of splits and offsets. The EVIPAS simulation-based optimization model, as well as other analytical design techniques, can recommend maximum green and gap setting values, but not coordination settings. As such, an additional analysis phase is sometimes recommended in the literature in order to obtain these settings. Skabardonis [1988], Courage [1989], Khatib and Coffelt [1999] have written guidelines and developed software packages associated with selection of coordination settings for actuated systems. The CORSIM Users Guide [ITT Systems and Sciences Inc., 1998] also contains guidelines for selection of coordination settings, though explained in the context of the CORSIM program.

Control Parameter Design and Optimization Summary

The control parameter design and optimization models in this subsection were presented primarily for informational purposes. In the context of this study, it is helpful to know that these models will not be immediate candidates for improvement of TRANSYT-7F performance, even though they exist in the literature as models for actuated control. Some of the concepts involving the individual traffic-actuated signal timing parameters were also discussed, which may be helpful as a review for the reader. Also, if it is ever decided that TRANSYT-7F could be enhanced in order to design or recommend traffic-actuated control parameter values, the models mentioned in this subsection may prove useful.

Chapter Two Summary

Many of the realistic components within these models for traffic-actuated control are not taken into consideration within the current methodology of TRANSYT-7F. However, not all of the available models in the literature for actuated control are able to meet the minimum requirements for improvement of TRANSYT-7F performance.

The following models do not meet the minimum requirements:

- Existing Model within TRANSYT-7F
- CORSIM
- EVIPAS
- Husch's Queue Service Time and Percentile Models
- Vehicle Delay Estimation Models
- Control Parameter Design and Optimization Models

The existing model within TRANSYT-7F is not ideal because it is oversimplified, which compromises the accuracy of actuated phase times. CORSIM and EVIPAS provide good treatment of actuated control, but from a practical standpoint, they are not good candidates for implementation in conjunction with TRANSYT-7F. Husch's queue service time model is essentially identical to the NCHRP queue service time model, and the percentile computation of actuated phase times is not ideal for internal links where random arrivals cannot be assumed. Finally, all vehicle delay estimation models, and all control parameter design and optimization models, do not meet the minimum requirement for this study associated with improved accuracy in calculating actuated phase times.

The following models do meet the minimum requirements for improvement of TRANSYT-7F performance:

- NCHRP Model
- Iterative Target Degree of Saturation Model
- Joint Poisson Probability Green Extension Time Model

These models were examined and tested further, and these results are documented in the upcoming chapters.

CHAPTER 3
MODEL DEVELOPMENT

This chapter describes the development of new models during the course of this study. An experimental version of TRANSYT-7F was developed in order to implement the new models, which address some of the existing deficiencies in estimating actuated phase times. Specific shortcomings of the target degree of saturation strategy, employed by multiple models from the literature, were discussed at length in chapter 2. Therefore, an ideal model for improvement of TRANSYT-7F performance should not adopt this strategy.

The queue service time — green extension time strategy, employed by multiple models from the literature, seems to be the most viable strategy. Thanks to the queue accumulation polygon (QAP) concept, which is well-documented within the Highway Capacity Manual and other sources, queue service time is readily and accurately calculated. Indeed, an actuated phase is not capable of gapping out in the midst of the queue service time, so this structure prevents modeling blunders. In addition, if the queue service time exceeds the maximum green time, max-out is the obvious result. Finally, taking queue service time out of the mix allows for more

accurate estimation of green extension time. Therefore, the queue service time — green extension time strategy will be pursued at this time.

Queue Service Time

In the literature, queue service time is computed by equations that are algebraic interpretations of the queue accumulation polygon. As mentioned in chapter 2, the QAP concept is flexible enough to adapt to complexities from the field that would alter vehicle arrival or departure rates. Here again is the queue service time equation (2-1) from the NCHRP model, initially introduced in chapter 2:

$$g_s = f_q \frac{q_r r}{(s - q_s)} \tag{2-1}$$

where,

q_r , q_g = red arrival rate (veh/s) and green arrival rate, veh/s, respectively,
r = effective red time, s,
s = saturation flow rate, veh/s, and
f_q = queue calibration factor

$$f_q = 1.08 - 0.1\left(\frac{\text{actual green}}{\text{maximum green}}\right) \tag{2-2}$$

Permitted Left-Turn Effects

One advantage of the queue service time approach involves the modeling of permitted left-turns. Typically permitted-only left-turn phases, in which only a green ball is displayed by the signal in the field, are not actuated because the length of the phase is designed to handle through movement traffic. However, protected-permitted left-turn

phases, in which a green arrow followed by a solid green is displayed in the field, are frequently actuated.

It is more complicated to estimate actuated phase times for the protected portion of a protected-permitted phase. The results are dependent on how much traffic is served during the permitted portion of the phase. If the opposing movement has heavy traffic and few left-turns are served during the permitted portion of the phase, then the actuated phase time may be nearly equal to what it would be under protected-only phasing. If the opposing movement has light traffic and many left-turns are served during the permitted portion, then the actuated phase time may be nearly equal to the minimum phase time, because the queues are so small during the protected portion. The queue accumulation polygon (QAP) illustrated in figure 3-1 shows that the queue length accumulated during the permitted portion directly affects the subsequent queue service time.

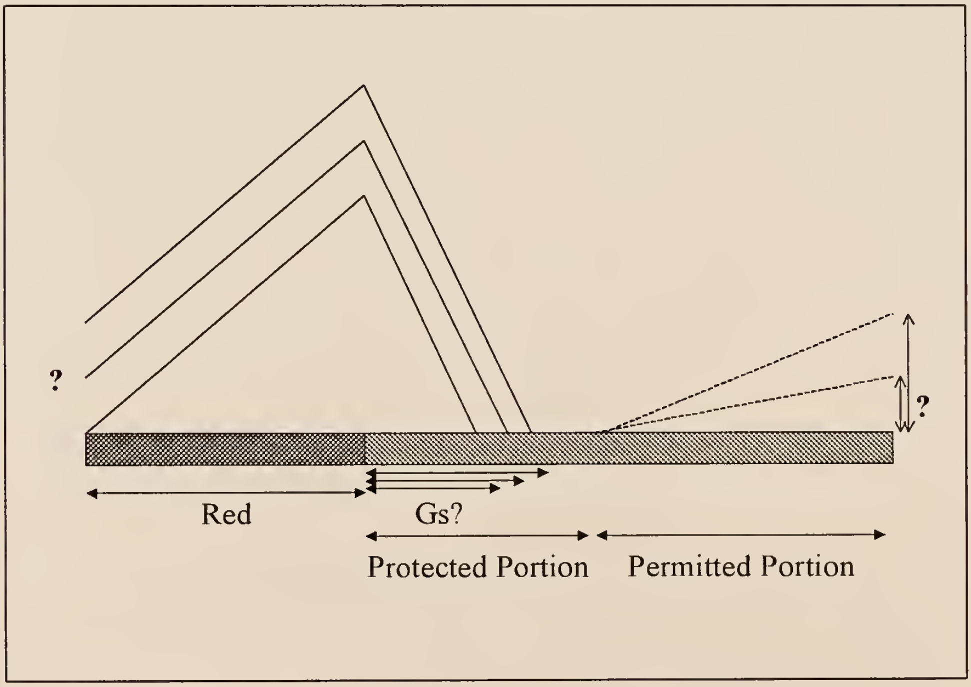


Figure 3-1: Queue Accumulation Polygon (QAP) for Protected-Permitted Phasing

The queue service time models are appropriately sensitive to how many vehicles were served during the permitted portion of the phase. Multiple QAPs from the HCM guidelines illustrate how the queue at the beginning of the protected portion of the phase can be adjusted to reflect the number of vehicles served during the permitted portion of the phase.

Whereas deficiencies of the existing literature models were initially introduced in chapter 2, additional technical details regarding these deficiencies are supplied within chapters 3 and 4. For example, chapter 2 states that permitted left-turn effects (on actuated phase times) are one category of operational characteristic effects that are difficult to quantify without using simulation. The existing model within TRANSYT-7F for calculating actuated phase times is inadequate because it does not attempt to estimate protected portion actuated phase times for protected-permitted phases. Instead such phases are currently set to their minimum phase times automatically, as if all traffic was served during the permitted portion of the phase. No known methodology is available for automatically adjusting the (protected portion) actuated phase target degree saturation in response to how much traffic was served during the permitted portion of the phase.

In order to rectify this model in its existing form, it would be necessary to readjust target degrees of saturation, based on the amount of traffic served during the permitted portion of the phase. Because phase times must be known prior to simulation, and because the amount of traffic served in the permitted portion is determined during simulation, an iterative target degree of saturation structure is required, similar to Akcelik's model from chapter 2.

If permitted left-turn effects were the only complexity affecting the actuated phase model, then the target degree of saturation strategy could possibly be updated as described by Akcelik. Perhaps there would be no obvious disadvantages relative to the queue service time — green extension time models; however, there are other complexities that further hinder this strategy, such as progression and spillback.

Progression and Spillback Effects

Unfortunately, the equations that interpret the QAP to compute queue service time are not flexible enough. They are oversimplified, and unable to interpret complex QAPs that are likely to occur due to network-wide interaction effects. The NCHRP queue service time equation implements two separate arrival rates, and one unique departure rate. The arrival rates are divided into two quantities: arrivals on red, and arrivals on green. The departure rate is the saturation flow rate. However, when analyzing a single system having multiple intersections, arrival rates and departure rates are unpredictable enough such that three variables are not adequate for calculating the correct results.

For example, consider having only one variable to describe arrivals on green. This means that only one arrival rate on green can be modeled. However, non-uniform arrival rates on green will occur when a nearby upstream signal is present. Even if the average arrival rate on green is specified correctly, progression effects (non-uniform arrivals) can change the queue service time.

Figure 3-2 illustrates an example of progression effects on queue service time. The slope of the queue accumulation polygon's right side is affected by a constantly changing arrival rate. Although 200 vehicles per hour are expected during the green phase, good progression from the upstream signal causes vehicles to arrive after the

queue has been serviced. In this case, progression effects allow the average queue service time to decrease from 12.6 seconds to 10.6 seconds. However, the formula would still predict a queue service time of 12.6 seconds in this case. One variable is not enough to account for progression effects. Ideally, the arrival rate during each second should be known in order to calculate the correct queue service time.

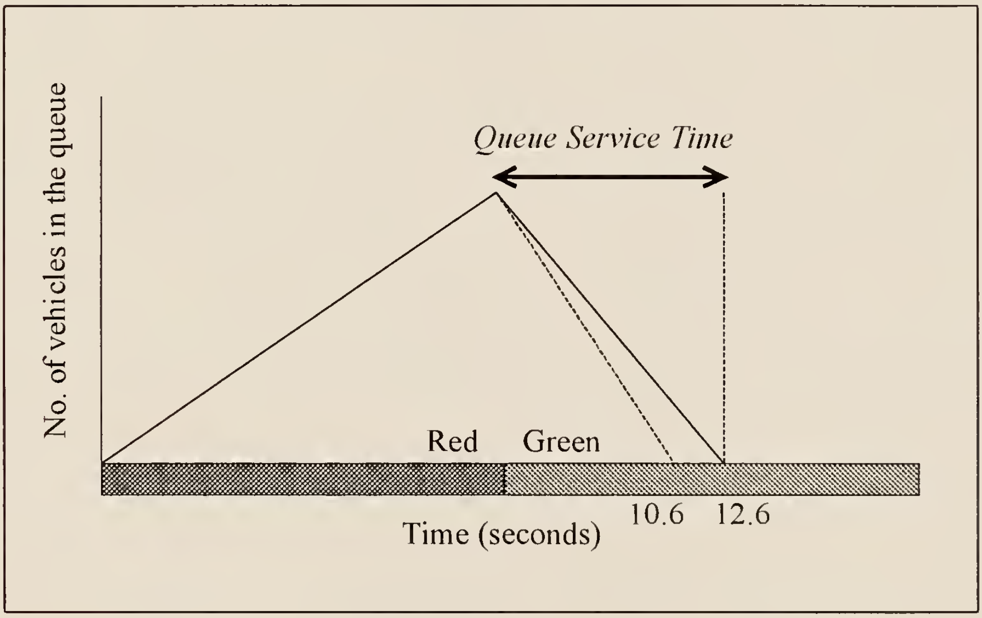


Figure 3-2: QAP Depiction of Good Progression Decreasing the Queue Service Time

The formula is additionally constrained to having only one variable to describe departure rates. This means that only one departure rate can be modeled for each movement. However, non-uniform departure rates will sometimes occur when a nearby downstream signal is present. Even if the average departure rate is specified correctly, spillback effects (non-uniform departure rates) can change the queue service time.

Figure 3-3 illustrates an example of spillback effects on queue service time. The slope of the queue accumulation polygon's right side is affected by a constantly changing departure rate. Although 3515 vehicle departures per hour are expected when serving queues, spillback from the downstream signal decreases the departure rate and increases the queue service time. In this case, spillback effects allow the average queue service time to increase from 15.6 seconds to 24.4 seconds. However, the formula would still predict a queue service time of 15.6 seconds in this case. One variable is not enough to account for spillback effects. Ideally, the departure rate during each second should be known in order to calculate the correct queue service time.

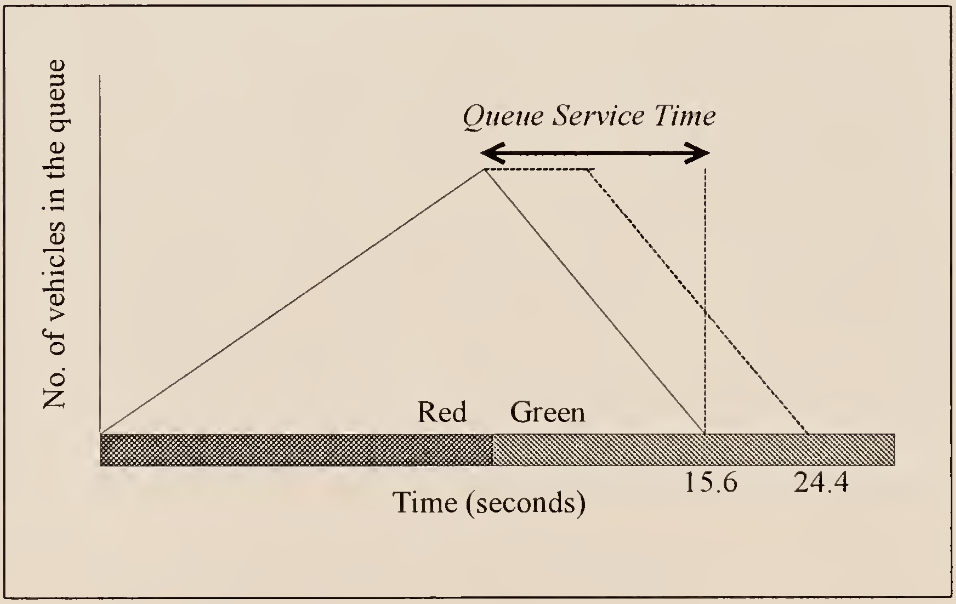


Figure 3-3: QAP Depiction of Spillback Increasing the Queue Service Time

Queue Service Times from the Program

Fortunately, TRANSYT-7F continuously tabulates uniform or non-uniform arrival rates, departure rates, and queue lengths during each step of analysis and

throughout the network. On external links with no nearby upstream signal, queue service times computed by TRANSYT-7F and the NCHRP formula are identical, barring any differences introduced by permitted left-turn movement models. However, TRANSYT-7F has the added capability to adjust queue service times in response to traffic flow from upstream signals, and spillback from downstream signals. Queue service time should therefore be extracted directly from the results of TRANSYT-7F step-wise simulation. This will allow queue service times and actuated phase times to be automatically responsive to progression and spillback effects, in addition to permitted left-turn effects.

Green Extension Time

A perceived deficiency of existing models in the literature for computing green extension time is the assumption of random vehicle arrivals. It is true that the assumption of random arrivals is reasonable on external links having no upstream signal nearby. However, many practitioners continue to conduct analyses with the assumption of uniform arrivals on external links. More importantly, an ideal green extension time model would not assume any specific pattern of vehicle arrivals, especially since platooned arrivals are expected on internal links having a nearby upstream signal. Rather, an ideal model would intelligently compute green extension times as a function of any possible pattern of vehicle arrivals. Here again is the green extension time equation (2-3) from the NCHRP model, initially introduced in chapter 2:

$$g_e = \frac{e^{\lambda(e_0+t_0-\Delta)}}{\varphi q} - \frac{1}{\lambda} \qquad (2\text{-}3)$$

where

e_0 = unit extension time setting

t_0 = time during which the detector is occupied by a passing vehicle

$$t_0 = \frac{(L_d + L_v)}{v} \tag{2-4}$$

where

L_v = vehicle length, assumed to be 5.5 m

L_d = detector length, DL, m,

v = vehicle approach speed, SP km/h

Δ = minimum arrival (intra-bunch) headway, s,

φ = proportion of free (unbunched) vehicles, and

λ = a parameter calculated as:

$$\lambda = \frac{\varphi q}{1 - \Delta q} \tag{2-5}$$

where q is the total arrival flow, veh/s for all lane groups that actuate the phase under consideration.

This equation noticeably implements one unique arrival rate (q). This means that only one arrival rate on green can be modeled; however, when analyzing a single system having multiple intersections, arrival rates are unpredictable enough such that one variable is not adequate for calculating the correct results. Non-uniform and non-random arrival rates on green will occur when a nearby upstream signal is present. Even if the average arrival rate on green is specified correctly, progression effects (non-uniform, non-random arrival rates) can change the green extension time.

Figure 3-4 illustrates an example of progression effects on green extension time. The histogram represents a constantly changing arrival rate. Although 200 vehicles per hour are expected during the green phase, good progression from the upstream signal causes vehicles to arrive after the queue has been serviced. In this case, progression

effects cause the average green extension time to increase from approximately 3 seconds to more than 5 seconds. However, the formula would still predict the same green extension time in both cases by assuming the average arrival rate of 200 vehicles per hour (illustrated by the superimposed dashed line). One variable is not enough to account for progression effects. Ideally, the arrival rate during each second should be known in order to calculate the correct green extension time.

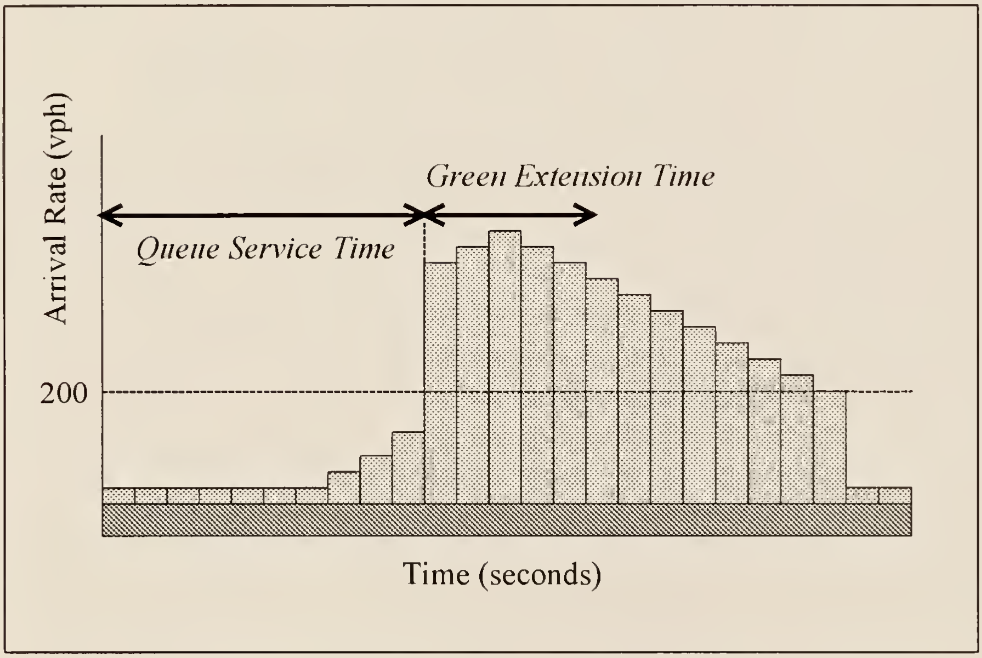


Figure 3-4: Example of Good Progression Increasing the Green Extension Time

Prototype Green Extension Time Model Using Applied Probability and Flow Profile

As mentioned earlier in this chapter, the TRANSYT-7F program continuously tabulates the arrival rate, departure rate, and queue length during each step of simulation and throughout the network. Knowledge of the network-wide arrival rate, or flow profile, should be useful in computing green extension time. However, the flow profile does not

automatically reveal the correct green extension time, in the way that queue profiles reveal the correct queue service time. The only way to explicitly compute green extension time is through microscopic simulation employed by programs like CORSIM. These programs simulate individual vehicles rolling over detectors and thus are able to know exactly when gap-out occurs. The TRANSYT-7F flow profile information allows the opportunity to calculate the maximum likelihood green extension time, which should be a better estimate than is possible using any model that assumes uniform or random arrivals. A new model was developed to compute the most probable location of gap-out and thus the overall green extension time, based on the TRANSYT-7F flow profile.

Methodology and sample calculation #1

A sample calculation illustrates the methodology of this prototype model. Figure 3-5 illustrates an actual simulated flow profile for an actuated left-turn phase having a nearby upstream signal. Although the average flow rate is 200 vehicles per hour, arrival rates much higher than this are occurring during a certain part of the 100-second cycle, due to progression effects. Note that flow intensity would be affected by multiple links in the case of a shared lane. Although figure 3-5 does not show this, the left-turn phase begins at step 92 towards the right-hand side of the histogram in this example. The force-off time for this phase occurs at step 29 towards the left-hand side of the histogram.

In order to estimate the green extension time, it is only necessary to scrutinize the flow profile after the queue service time and until the force-off time. Figure 3-6 illustrates this abbreviated flow profile that can be used for green extension time calculations. Looking at this flow profile, the flow rate occurring immediately after queue discharge is clearly higher than the 200 vph average. In fact, the average flow rate

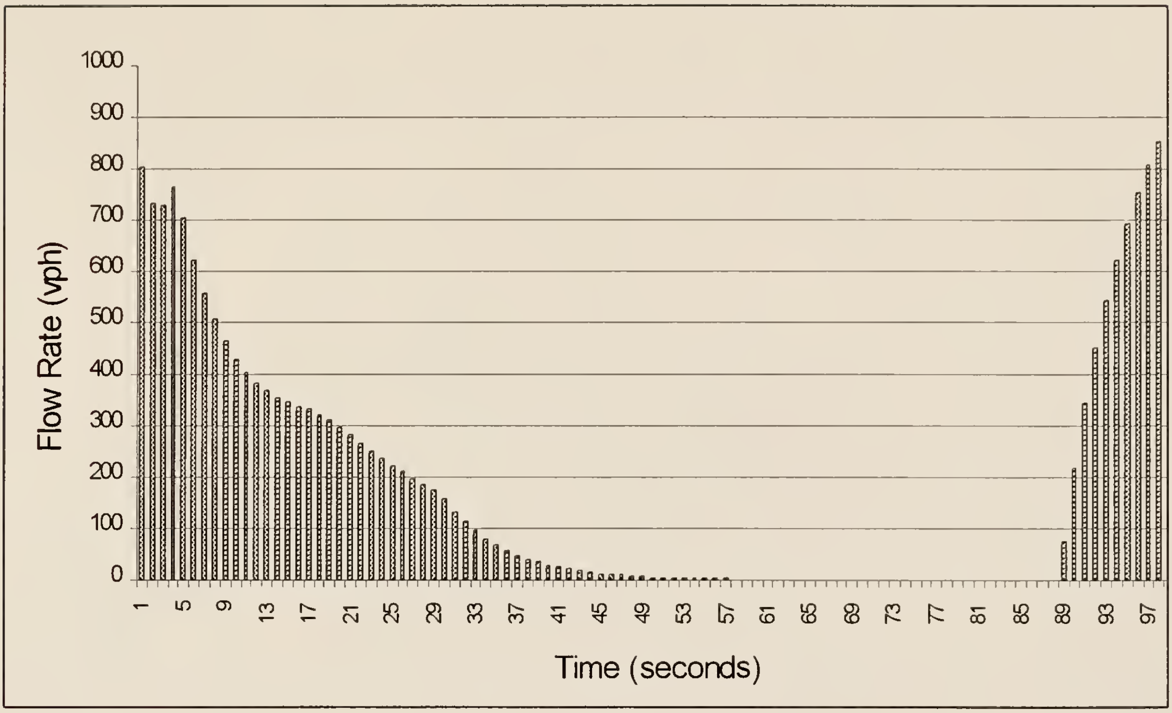


Figure 3-5: Sample Calculation Internal Link Flow Profile

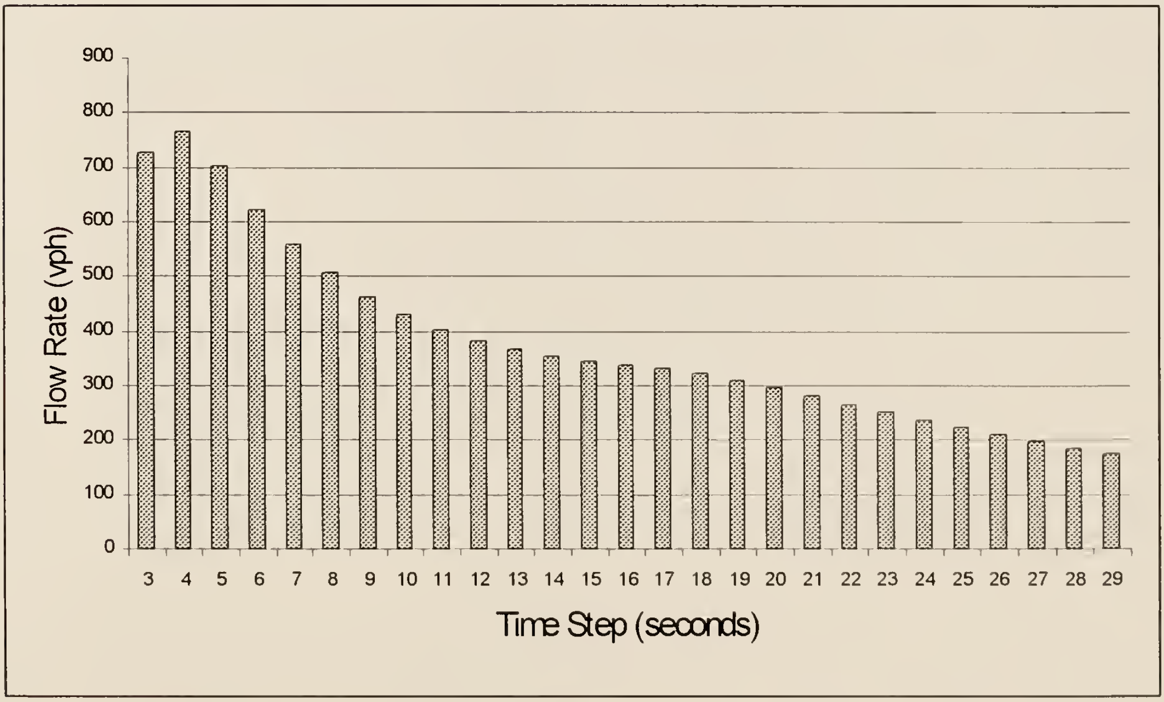


Figure 3-6: Sample Calculation Internal Link Flow Profile (Abbreviated)

in this range is 380 vph. Therefore, a higher-than-average green extension time is expected.

The first step of the prototype model involves calculating the probability of zero vehicle arrivals during each second within the abbreviated flow profile. The Poisson distribution is used to calculate the probability of zero arrivals during each second. However, instead of using the average hourly traffic volume in the calculations, each individual flow rate at each individual simulation step is used to calculate numerous individual probabilities. For example, the probability of zero arrivals at step 3 immediately following queue discharge, is:

$$P(0) = e^{-727/3600} \approx 0.82$$

The second step of applying the prototype model involves calculating the probability of consecutive simulation steps having zero vehicle arrivals within the abbreviated flow profile. It is known that if zero arrivals take place on enough consecutive steps, the phase will terminate via gap-out. Assuming a gap setting of 3 seconds, phase termination will occur if three consecutive steps have zero arrivals. For example, the probability of three consecutive steps with zero arrivals at step 5 immediately following queue discharge, is:

$$P(0,0,0) = 0.82 \times 0.81 \times 0.82 \approx 0.54$$

The third step of applying the prototype model involves calculating the probability of gap-out within the abbreviated flow profile. At step 5, the probability of gap-out is actually equal to the probability of consecutive zero arrivals at steps 3, 4, and 5. However, at step 8, the probability of gap-out is not exactly equal to the probability of

consecutive zero arrivals at steps 6, 7, and 8. It is known that there is only a 1 - 0.54 = 0.46 probability of not gapping out at step 5. Therefore, the probability of gap-out at step 8 is:

$$P(\text{gap-out}) = 0.46 \times (0.84 \times 0.86 \times 0.87) \approx 0.29$$

The fourth step of applying the prototype model involves calculating the probability of max-out within the abbreviated flow profile. The probability of max-out is actually equal to one minus the summation of all gap-out probabilities:

$$P(\text{max-out}) = 1 - \sum P(\text{gap-out})$$

The fifth step of applying the prototype model involves calculating time step weighted averages. In this sample calculation, there was a 0.54 probability of phase termination at step 5. Therefore, the time step weighted average here is:

$$T(\text{step } 5) = 0.54 \times 5 \approx 2.7$$

The sixth step of applying the prototype model involves calculating the average phase time by summing the time step weighted averages:

$$PhaseTime = \sum T(\text{step})$$

Table 3-1 lists the entire set of values in the sample calculation. After summing the time step weighted averages, the final answer is time step 7.1 as the location of phase termination. Because the queue service time ended at step 2, the green extension time is 5.1 seconds. The Poisson distribution is not the only method available to calculate $P(0)$. For 200 vph (an arrival every 18 seconds on average), the uniform $P(0)$ would be 1/18, or

0.055. However, this method would only change the answer by one tenth of a second in table 3-1. Note that the probability of immediate gap-out was 54%. Oddly, Husch's green extension time model specifies immediate gap-out due to the value above 50%, even though the 46% chance of extension results in a 2.1 second increase (g_e = 5.1 vs. 3). Also note that green extension times are potentially affected by the maximum green setting. Green extension time decreases to 4.4 seconds if the maximum occurs at step 8, because all of the remaining weighted average values would be rolled into step 8.

Table 3-1: Prototype Model Sample Calculation of Green Extension Time

Time Step	Flow Rate (vph)	P(0)	P(0,0,0)	P(gap)	P cumulative	Weighted Average
3	727	0.82				
4	765	0.81				
5	702	0.82	0.54	0.540		2.700
6	623	0.84				
7	558	0.86				
8	506	0.87	0.63	0.288	0.83	2.303
9	464	0.88				
10	430	0.89				
11	402	0.89	0.70	0.120	0.95	1.321
12	383	0.90				
13	367	0.90				
14	354	0.91	0.74	0.038	0.99	0.536
15	345	0.91				
16	337	0.91				
17	331	0.91	0.75	0.010	1.00	0.176
18	323	0.91				
19	311	0.92				
20	297	0.92	0.77	0.003	1.00	0.052
21	282	0.92				
22	266	0.93				
23	251	0.93	0.80	0.001	1.00	0.014
24	237	0.94				
25	223	0.94				
26	211	0.94	0.83	0.000	1.00	0.003
27	198	0.95				
28	186	0.95				
29	175	0.95		0.000		0.001
						7.107

Another observation is that green extension time calculations begin at an integer step number such as step 3 because TRANSYT-7F tabulates the flow profile by integer step sizes. Although queue service time is actually a real number and the green extension time actually begins at time step 2.6, the approximation of beginning the green extension time calculations at step 3 should have negligible impact on the result (5.1 seconds). In other words, there are so many calculations within the flow profile, that beginning the green extension time calculations at step 2 or step 3 would probably produce the same answer (5.1 seconds).

Sample calculation #2

In order to better understand the prototype model, another sample calculation can be demonstrated in which changes to progression cause changes in the green extension time. In this sample calculation, all conditions except for one are identical to those from the previous sample calculation. The one difference is the offset design, which changes the observed traffic patterns on the arterial street. Figure 3-7 illustrates the updated flow profile for the actuated left-turn phase having a nearby upstream signal.

Comparing the figure 3-7 flow profile with the original flow profile from figure 3-5, the shape and intensity of the platoon is similar. However, with the new offset design, the platoon arrives 5 seconds earlier at step 84. Again, to estimate the green extension time, it is only necessary to scrutinize the flow profile after the queue service time and until the force-off time. Figure 3-8 illustrates the updated abbreviated flow profile that can be used for green extension time calculations.

The queue service time ends at step 12, and the force-off time occurs at step 37. Looking at this flow profile, the flow rate occurring immediately after queue discharge is

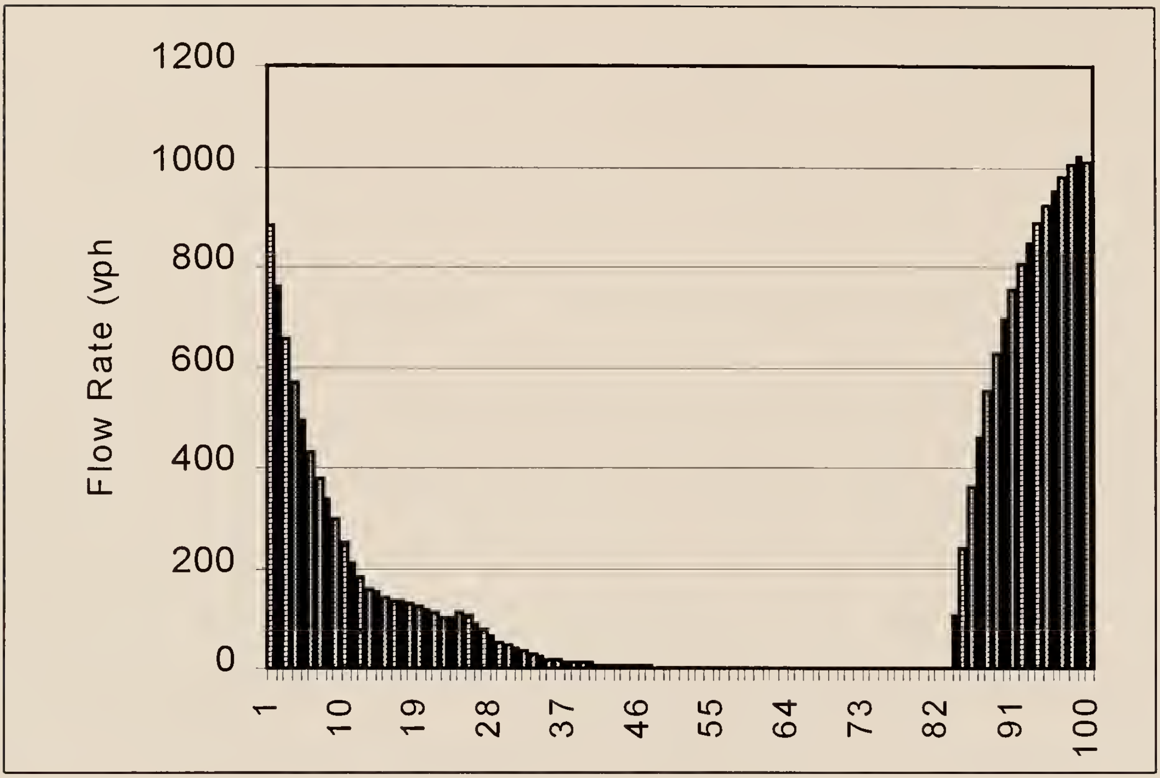


Figure 3-7: Sample Calculation #2 Internal Link Flow Profile

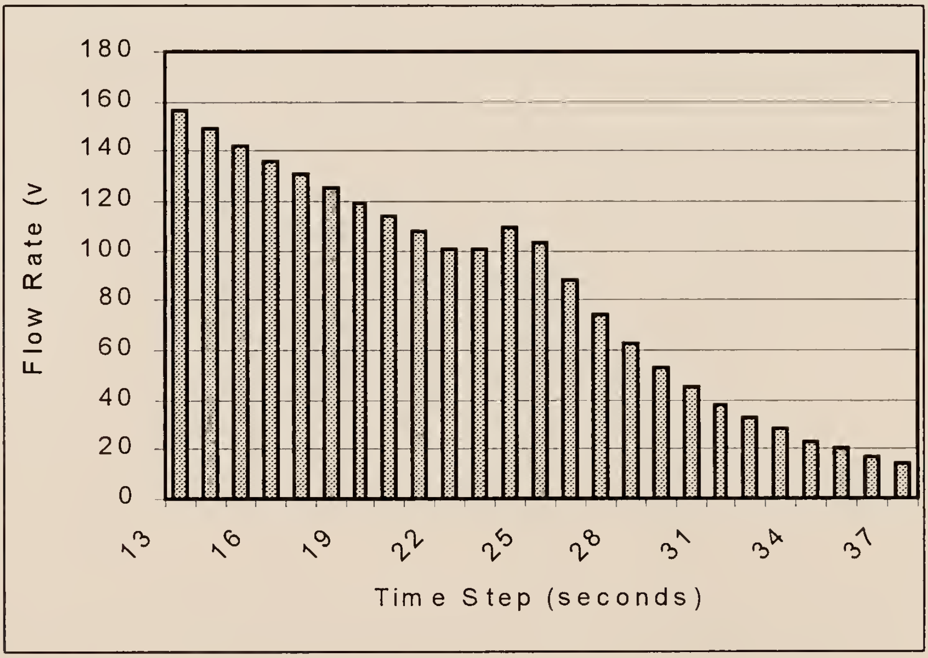


Figure 3-8: Sample Calculation Internal Link Flow Profile (Abbreviated)

clearly lower than the 200 vph average. In fact, the average flow rate in this range is 84 vph. Therefore, a lower-than-average green extension time is expected. Table 3-2 lists the entire set of values in the second sample calculation.

Table 3-2: Prototype Model Sample Calculation #2 of Green Extension Time

Time Step	Flow Rate (vph)	P(0)	P(0,0,0)	P(gap)	P cumulative	Weighted Average
13	156	0.96				
14	149	0.96				
15	142	0.96	0.88	0.88		13.200
16	136	0.96				
17	131	0.96				
18	125	0.97	0.90	0.108	0.99	1.937
19	119	0.97				
20	114	0.97				
21	108	0.97	0.91	0.011	1.00	0.236
22	101	0.97				
23	101	0.97				
24	109	0.97	0.92	0.001	1.00	0.025
25	103	0.97				
26	88	0.98				
27	74	0.98	0.93	0.000	1.00	0.002
28	63	0.98				
29	53	0.99				
30	45	0.99	0.96	0.000	1.00	0.000
31	38	0.99				
32	33	0.99				
33	28	0.99	0.97	0.000	1.00	0.000
34	23	0.99				
35	20	0.99				
36	17	1.00	0.98	0.000	1.00	0.000
37	14	1.00				0.000
						15.401

After summing the time step weighted averages, the final answer is time step 15.4 as the location of phase termination. Because the queue service time ended at step 12, the green extension time is 3.4 seconds. The sample calculation demonstrates that when progression becomes worse, the prototype model responds logically by calculating a

lower green extension time. This means that because the traffic flow pattern is light following queue discharge, there is a lower probability of the phase being extended. By comparison, the NCHRP model would not have recognized progression effects. It would have calculated the same green extension time in both scenarios.

Prototype Green Extension Time Model Based on Uniform Vehicle Arrivals

Another prototype model was developed during the course of this study as an alternative to the existing models from the literature designed with random arrivals in mind. This model carries with it the basic assumption of uniform vehicle arrivals. The model is simple because when vehicle arrivals may be assumed to be perfectly uniform, there are only two green extension times that are physically possible. After computing a weighted average based on these two possible green extension times, the result is an overall estimate for the average green extension time.

When perfectly uniform arrivals are in effect, the two physically possible green extension times can be described as follows. The first possibility is that no vehicle arrives during the gap setting interval immediately following queue discharge. When this occurs, the green extension time is equal to the gap setting. The second possibility is that one vehicle arrives during the gap setting interval immediately following queue discharge. When this occurs, the green extension time is equal to 1.5 times the gap setting. If a vehicle extends the phase once, no additional extensions are possible because the inter-arrival time between vehicles will be larger than the gap setting. Assuming that one vehicle does extend the phase, it must arrive sometime within the gap setting interval, or on average halfway through the gap interval. After that, the phase will be extended by exactly 1 additional gap time.

Sample calculation

A sample calculation is useful for understanding the logic here. Suppose that the traffic volume demand is 600 vehicles per hour per lane. Assuming uniform arrivals, this translates into one vehicle arrival every six seconds. Given a gap setting of two seconds, there would be a 0.33 probability of a vehicle arrival within the gap interval, and a 0.67 probability of no vehicle arrival within the gap interval. Also, if a vehicle does arrive within the gap interval, then on average it will arrive halfway through the gap interval, or after 1 second in this case. In addition, a vehicle extension after 1 second automatically extends the phase by exactly 1 gap interval, or 2 seconds in this case. Therefore, there would be a 0.33 probability that the green extension time would be equal to 1.5 * 2 = 3 seconds, and a 0.67 probability that the green extension time would be equal to 2 seconds. Using the weighted average, overall green extension time is estimated as:

$$g_e = (0.33 \times 3) + (0.67 \times 2) = 2.33 \text{ seconds}$$

These green extension time calculations are only valid given traffic volumes that are sufficiently low. High traffic volumes result in low inter-arrival times that may be lower than the gap interval. When vehicle inter-arrival times are indeed lower than the gap interval, and uniform arrivals are in effect, the end result is that gap-out cannot occur, physically or mathematically. In this case, the phase is assumed to terminate via max-out or force-off. For example, a per lane volume of 1800 vehicles per hour, producing one vehicle arrival every two seconds, would always cause max-out with gap settings of two seconds or higher. Likewise, a per lane volume of 1200 vehicles per hour, producing one

arrival every three seconds, would always cause max-out with gap settings of three seconds or higher.

Alternative queue service time model

Interestingly, this second prototype green extension time model concept suggests a new and simplified calculation for queue service times. Indeed, if vehicle arrivals are perfectly uniform in nature, and if only 0-1 vehicle extensions of the green are physically possible, this implies that nearly all vehicles are queued while being served. Thus, the uniform arrivals queue service time could presumably be calculated by computing the number of vehicles served per cycle, multiplying this value by the queue discharge headway, and then adding in the start-up lost time. Figure 3-9 illustrates the phase time calculation under this scenario.

Figure 3-9 shows that four queued vehicles should be served on each cycle, although the fourth vehicle joins the queue in the midst of queue service time. The phase time of 15 seconds assumes no vehicle extensions. However, since the uniform arrivals green extension time should actually be the weighted average of 0 and 1 extension times, the estimated phase time would probably be something like 15.75 seconds.

When the uniform queue service times are combined with the uniform green extension times, results may improve due to the consistency of assumptions. However, the process of determining an average number of vehicles served per cycle, which allows computation of uniform queue service times, is fragile. Determination of average queue per cycle can become difficult, given numerous complications from the field, which is why extraction of queue service times from TRANSYT-7F simulation is preferable. Consequently, the uniform arrivals prototype model is expected to produce reliable

results under simple conditions, but may break down under complex conditions. In order to adapt the model to complex conditions, it would be necessary to somehow adjust the expected average queue per cycle in response to numerous factors (permitted left-turns, spillback, etc.).

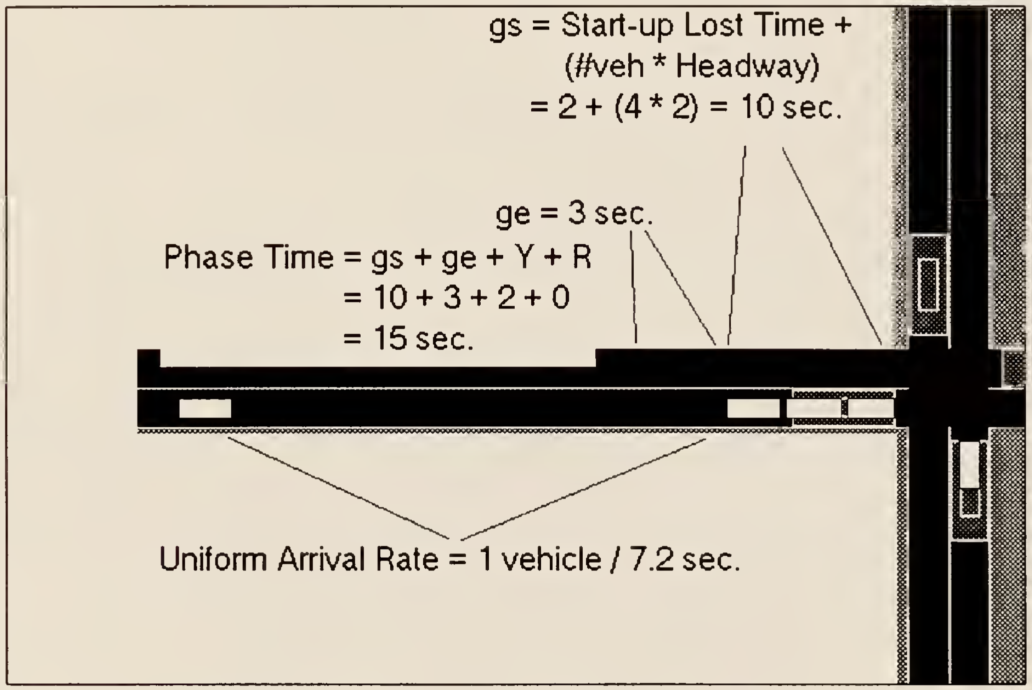


Figure 3-9: Calculating Queue Service Times with the Uniform Arrivals Assumption

The uniform arrivals prototype model may provide a good reference point during the testing process, and could be relatively effective at isolated intersections, where conditions tend to be less complicated. Practical advantages include better computing speed and easy programming, since it is not necessary to obtain queue or flow profile data from the optimization program.

Preliminary Model Comparison

Thorough testing of all candidate models is documented in chapter 4. For the sake of understanding the two prototype models for computing green extension times, some preliminary comparison is useful.

Model comparison under platooned arrivals

Recall that the sample calculations for the first prototype model involved an actuated, major street left-turn phase having platooned vehicle arrivals due to a nearby upstream signal. The traffic volume was 200 vph with a gap setting of 3 seconds. The first prototype model had predicted green extension times of 5.1 and 3.4 seconds, depending on the offsets and progression on the arterial street. Under these conditions, the second prototype model would calculate a green extension time based on one vehicle arrival every 18 seconds, and a one in six chance of an arrival during the gap interval:

$$g_e = (0.17 \times 11.5 \times 3) + (0.83 \times 3) = 3.25 \text{ seconds}$$

Note that this green extension time would be applicable in both sample calculations because the second prototype model does not take progression effects into account. In addition, regardless of progression effects, the NCHRP model computes a single green extension time (4.6 seconds), as does Husch's Poisson probability model (3 seconds). Depending on the duration of the queue service time, differences in predicted phase times will not necessarily be equal to the differences in green extension times. This will be illustrated by a larger set of results in chapter 4.

Model comparison under uniform arrivals

Another interesting comparison between the prototype green extension time models occurs on external links with no nearby upstream signal. On these links, vehicle arrivals tend to be nearly uniform in nature, or perhaps random arrivals with uniform inter-arrival times on average. In theory, the first prototype model should be able to calculate good results under this scenario because uniform arrivals are simply another variation on the flow profile. Certainly the second prototype model, designed with uniform arrivals in mind, would be expected to calculate relatively accurate green extension times on external links, where vehicle arrivals are known to be close to uniform in nature.

Suppose the green extension time must be calculated on an external link having 500 vehicles per hour, with a gap setting of 3 seconds. Figure 3-10 illustrates the uniform flow profile, and table 3-3 lists all calculations from the first prototype model. In this case the calculated green extension time is 4.5 seconds, because queue service time ends at step 12 and gap-out occurs at step 16.5. Once again, substitution of uniform probabilities in place of Poisson probabilities of zero arrivals at step would only change the final answer by one tenth of a second.

Although the first prototype model calculates a green extension time of 4.5 seconds, the NCHRP model and Husch's Poisson probability model calculate 6.1 and 3 seconds respectively for the same conditions. Since the 500 vph volume generates one arrival every 7.2 seconds, a 3-second gap results in a 3/7.2= 0.42 chance of phase extension, the second prototype model computes:

$$g_e = (0.42 \times 1.5 \times 3) + (0.58 \times 3) = 3.63 \text{ seconds}$$

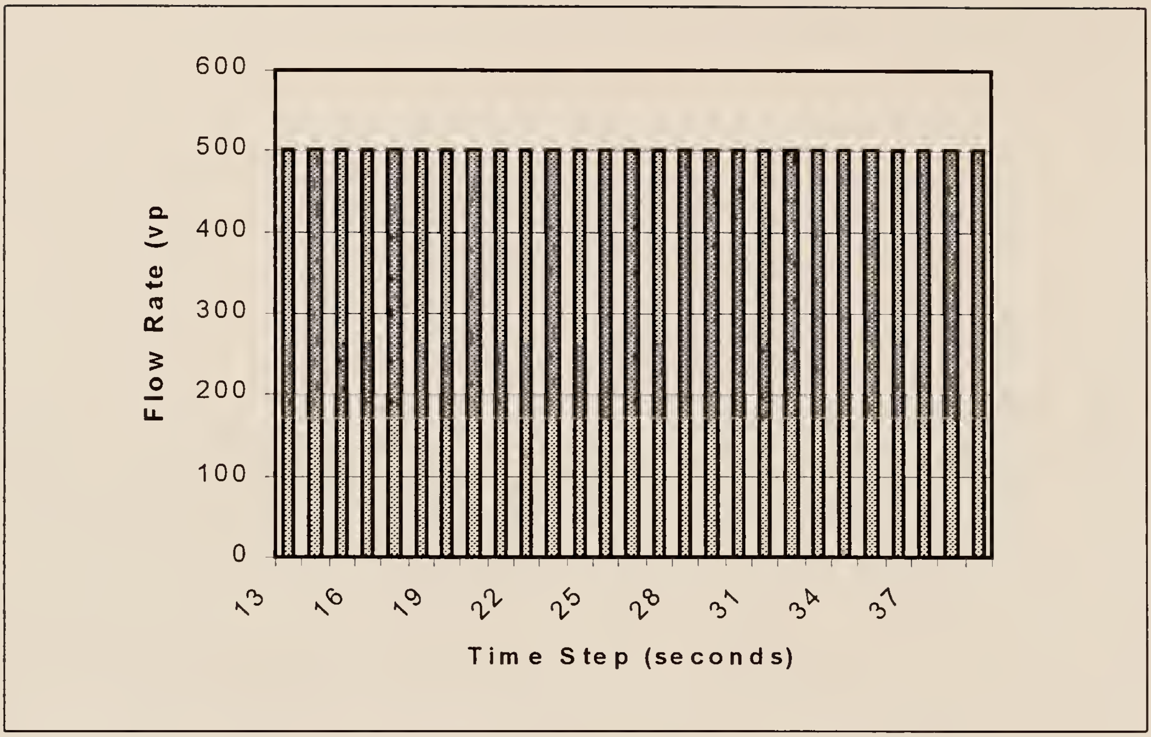


Figure 3-10: Sample Calculation External Link Flow Profile (Abbreviated)

Preliminary model comparison summary

The differences in results from the two prototype models, given uniform arrivals, indicate that the first prototype model may be overestimating green extension times. However, results from the NCHRP model indicate the opposite. In the absence of additional test results, the only observation to be made on the first prototype model at this time is that it responds appropriately to maximum green and progression effects. More testing results are necessary to determine the accuracy of each candidate model. Extensive testing results are presented in chapter 4.

Table 3-3: Calculations of the First Prototype Model Under Uniform Arrivals

Time Step	Flow Rate (vph)	P(0)	P(0,0,0)	P(gap)	P cumulative	Weighted Average
13	500	0.87				
14	500	0.87				
15	500	0.87	0.66	0.66		9.900
16	500	0.87				
17	500	0.87				
18	500	0.87	0.66	0.224	0.88	4.035
19	500	0.87				
20	500	0.87				
21	500	0.87	0.66	0.076	0.96	1.604
22	500	0.87				
23	500	0.87				
24	500	0.87	0.66	0.026	0.99	0.625
25	500	0.87				
26	500	0.87				
27	500	0.87	0.66	0.009	1.00	0.239
28	500	0.87				
29	500	0.87				
30	500	0.87	0.66	0.003	1.00	0.091
31	500	0.87				
32	500	0.87				
33	500	0.87	0.66	0.001	1.00	0.034
34	500	0.87				
35	500	0.87				
36	500	0.87	0.66	0.000	1.00	0.013
37	500	0.87				
38	500	0.87				
39	500	0.87	0.66	0.000	0.00	16.540

Model Implementation

Isolated Versus Coordinated Operation

Control type (isolated vs. coordinated) is expected to have a fundamental effect on actuated phase time results. The top half of figure 3-11 illustrates the phasing diagrams under coordinated conditions. Diagram #1 illustrates typical operation under

the conditions from the first part of experiment #1. Suppose that a certain parameter, such as detector length, free flow speed, or vehicle length, has its value increased such that green extension time would tend to respond by increasing as well (diagram #2). Although this change by itself would tend to increase the average phase time, a feedback effect exists that reduces the upcoming queue service time. This is because the effective red time for phase #1 is equal to the length of phase #2. When the length of phase #2 is reduced, this reduces the effective red time for phase #1, thus reducing the subsequent queue and queue service time for phase #1 (diagram #3).

The bottom half of figure 3-11 illustrates the phasing diagrams under isolated conditions. Diagram #1 illustrates typical operation under the conditions from the first part of experiment #1. Suppose that a certain parameter, such as detector length, free flow speed, or vehicle length, has its value increased such that green extension time would tend to respond by increasing as well (diagram #2). This change tends to increase not only the length of phase #1, but also the effective red time imposed on actuated phase #2. This in turn increases not only the queue, queue service time, and length of phase #2, but also the effective red time imposed on actuated phase #1 (diagram #3). At first glance this appears to be an infinite loop, but as described in chapter 2, this process converges reliably to a new actuated cycle length.

These results for isolated conditions are expected and documented in chapter 2 and the HCM guidelines. However, the results for coordinated conditions are interesting. They imply that certain parameters (e.g. detector length, approach speed, vehicle length, maximum green) which when increased would normally tend to increase the green extension time, perhaps do not significantly increase the phase time due to the queue

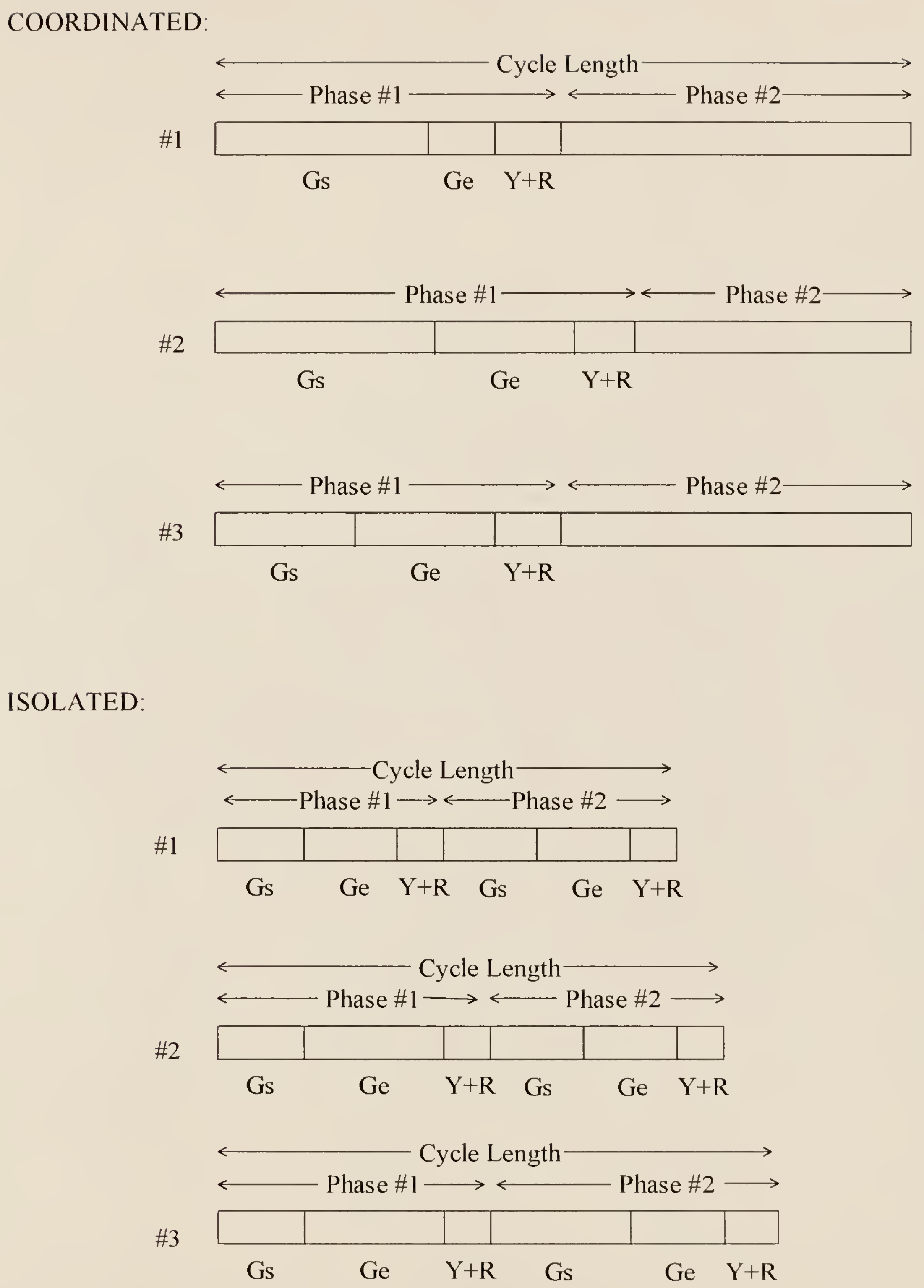


Figure 3-11: Feedback Effects under Coordinated and Isolated Conditions

service time feedback effect. This begs the question of whether some parameters affect the phase time at all. Do the queue service time and green extension time merely redistribute themselves such that the phase time is unmovable, or do phase time shifts occur anyway? This could be important because if certain parameters do not affect phase time, this simplifies the problem considerably and allows for shortcuts to be taken in modeling. Testing results related to this concept are discussed later on in chapter 4.

To summarize, fundamental differences in the actuated phase times are expected when comparing results under different signal control types (coordinated vs. isolated). These differences are likely caused by the queue service time — green extension time feedback effect. In the context of TRANSYT-7F and coordinated signal systems, correct understanding and modeling of coordination behavior is a higher priority. However, important analyses are performed on isolated intersections also. Proper understanding and modeling of these conditions is also useful.

Early Return to Green Effects

One of the first things observed when looking at preliminary results on arterial street performance was that some of the oversaturated actuated phase lengths were being underestimated by the TRANSYT-7F candidate models, relative to CORSIM. Figure 3-12 illustrates the pitfall that was occurring. CORSIM was showing that unused green time from the first actuated phase #1 was being taken by oversaturated phase #2. However, allocation of unused green time was originally been performed according to the NCHRP procedure, described in chapter 2, which specifies that all unused green time will be donated to the non-actuated phase under coordinated conditions.

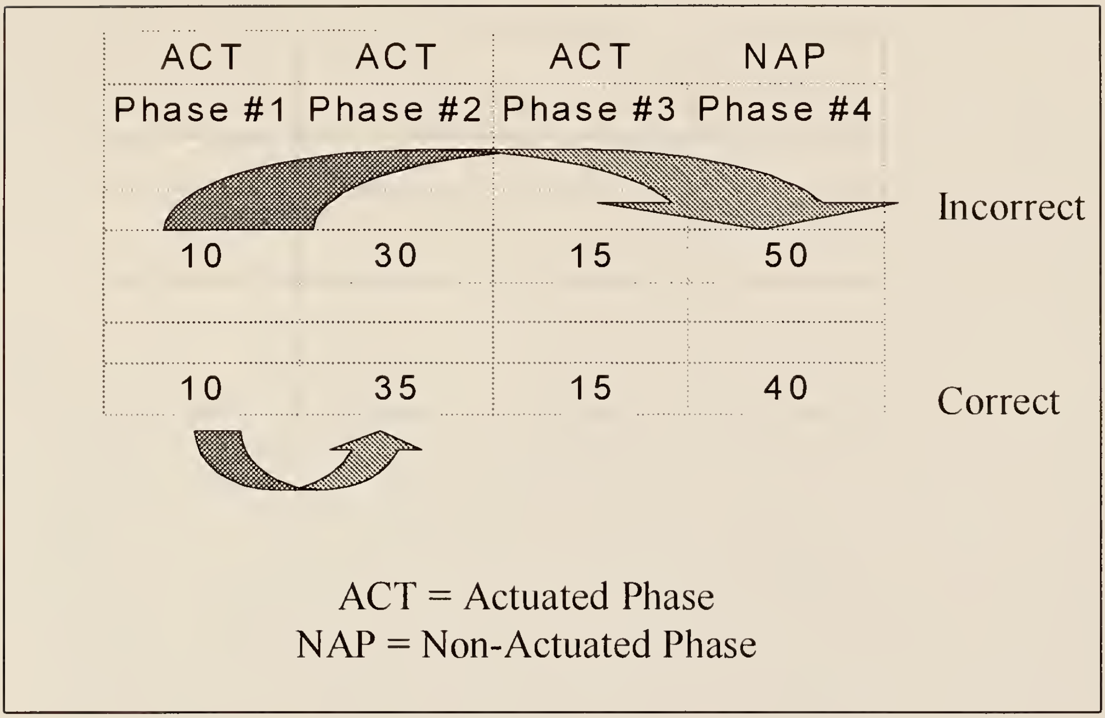


Figure 3-12: Allocation of Unused Green Time under Coordinated Conditions

Before any significant testing was performed on arterial streets, the experimental version of TRANSYT-7F was redesigned such that oversaturated actuated phases would be able to take unused green time from prior actuated phases if necessary. Note that this means actuated phase #1 from figure 3-12 is not able to take unused green time from the other actuated phases because they occur later on in the cycle. If no oversaturated actuated phase occurs following an undersaturated actuated phase, unused green time will be utilized by the non-actuated phase.

Overlap Phasing Effects

The next obvious problem observed when looking at preliminary results on arterial street performance was that some of the undersaturated actuated phase lengths were being overestimated by the TRANSYT-7F candidate models. This was occurring

due to a consistent, incorrect determination of the "critical link" within actuated phases. What is a critical link?

In experiment #1 only one link was moving during each phase. However, under complex conditions, there will be multiple links moving in each phase. When this happens, the basic strategy of an actuated controller dictates that the right-of-way should be terminated after all queues have been served on all links. Therefore, the overall phase should be terminated after the longest queue has been served, or rather after the queue that takes the longest time to dissipate has been served. The link that happens to possess relatively heavy traffic, resulting in a queue that takes the longest time to dissipate, is known as the critical link.

Figure 3-13 illustrates a hypothetical phase sequence where multiple links are moving on each phase. If the hourly traffic volume demand on each link is known, then normally the link having the highest volume during each phase would be critical and affect phase termination. For example, the link having 300 vph will affect when phase #1 terminates, the link having 500 vph will affect when phase #2 terminates, etc.

It may take more than knowledge of the traffic volume in order to determine the true critical link. Suppose the 500 vph link in phase #2 actually has two lanes to use, unlike the 400 vph link that only has one lane to use. This means the queues on the 500 vph link will dissipate twice as quickly, behaving almost like two 250 vph lanes, and the 400 vph link will then have the longest queues on average and tend to affect phase termination.

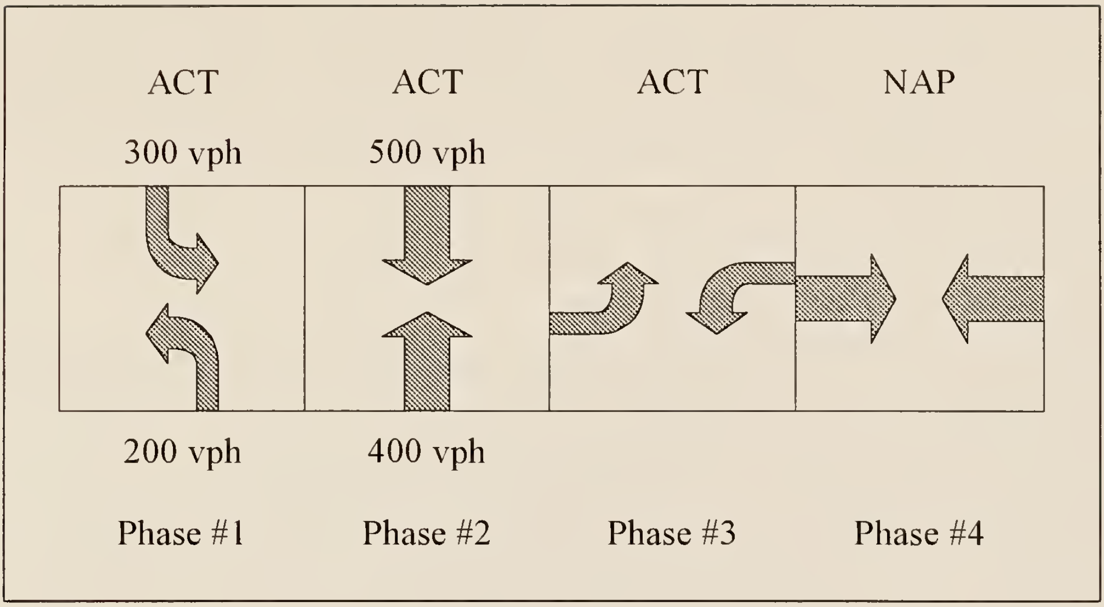


Figure 3-13: Sample Problem for Determination of Critical Link

In addition to number of lanes, there are numerous field conditions that can slow down or speed up queue service time, including vertical grade, lane width, heavy vehicle percentage, percentage of turns from a shared lane, and parking or bus maneuvers, just to name a few. The effect of these field conditions is quantified by the saturation flow rate (s) parameter.

The flow ratio (v/s) can be used to quantify the combined affect of volume and saturation flow rate. This becomes a better quantity to use than volume in determining the critical link because the queue service rate is taken into account. Thus, in the early stages of data collection for arterial streets, flow ratio was used to determine the critical links. Unfortunately, critical link determination using this parameter can be shown to be inadequate due to the overlap phasing effects frequently observed at actuated signals, as illustrated in figure 3-14.

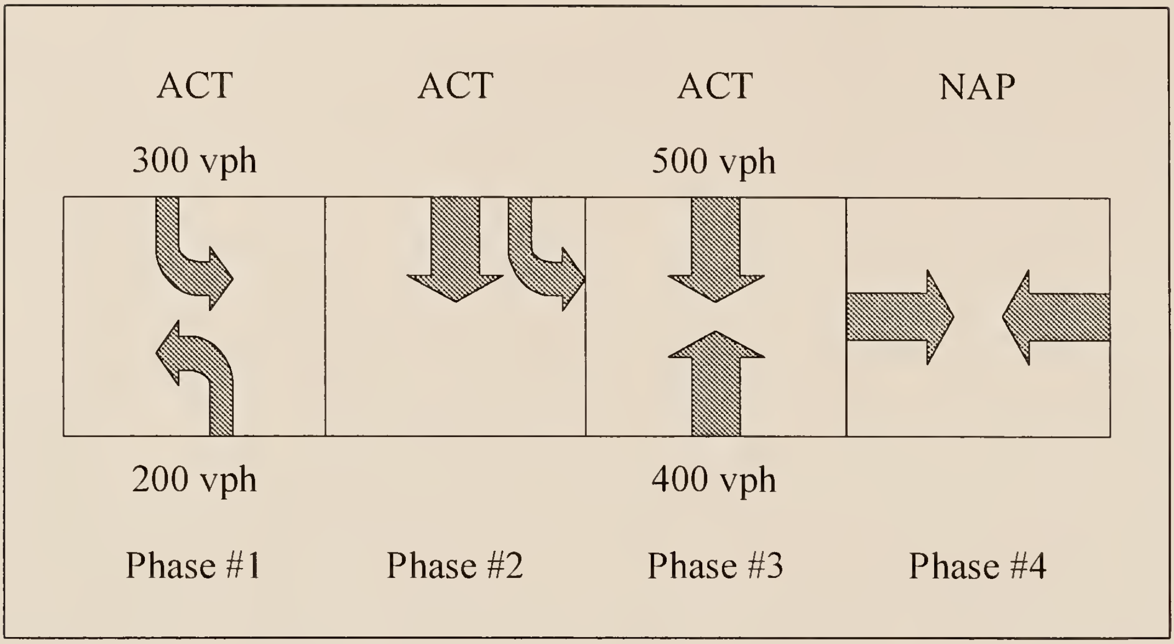


Figure 3-14: Sample Determination of Critical Link with Overlap Phasing Effects

Figure 3-14 shows that the 200 vph actuated left-turn phase terminates earlier than the 300 vph left-turn phase. This means that the 500 vph through movement link, which is adjacent to the 300 vph left-turn link, gets a head start and begins to move earlier (phase #2) than the opposing 400 vph through movement link. Assuming these links have the same saturation flow rate, this means that even though the 500 vph link has the highest flow ratio (v/s), its queue doesn't necessarily take longer (than the 400 vph link) to dissipate because its vehicles begin moving earlier in the cycle. If flow ratio was the only technique available for determining critical link, the non-critical link would sometimes be chosen as critical by the candidate models, resulting in underestimated phase times.

Because of the overlap phasing effects, the only way to really know which link is critical and will extend the phase is to apply the candidate models to each candidate link

that may potentially be critical, or to perform what could be called a critical link search. This means that for each link it is necessary to compute queue service time and green extension time, and then to compare all phase times to determine which one would terminate later in the cycle. This significantly increases the program running time, but is necessary in order to obtain accurate and realistic actuated phase times. Therefore, before any significant testing was performed on arterial streets, the experimental version of TRANSYT-7F was redesigned such that it would by default perform a critical link search for each actuated phase, instead of using the flow ratio to determine critical links.

Competing Link Effects

After updating the experimental version of TRANSYT-7F such that correct allocation of unused green time and correct critical link determination would occur, phase times were again compiled for an arterial street test data set. Table 3-4 contains the actuated phase times at each of the 10 intersections from one of the TRANSYT-7F candidate models (top) and CORSIM (bottom). This table contains shaded cells to indicate phases having multiple actuated links that could potentially be critical links. In this experiment, average phase times from CORSIM were calculated based on individual phase times observed over approximately 30 minutes of simulation (18 simulated cycles).

One observation available from the table 3-4 results is that the candidate model is underestimating a couple of the phase times in shaded boxes where multiple actuated links are moving. This occurs even though the program had already been updated to perform critical link search and correctly locate the critical link as mentioned earlier. The underestimated phase times from the candidate model are noticeable at intersection 2

Table 3-4: Examples of Competing Link Effects on Actuated Phase Times

	Intersection #	Phase #1	Phase #2	Phase #3	Phase #4	Phase #5
TRANSYT-7F Candidate Model	1	13	19	15	30	
	2	12	18	15	18	
	3	17	17			
	4	13	25			
	5	12	18	12	15	16
	6	13	18	12	15	30
	7	12	15	15	30	
	8	12	17	15	30	
	9	12	18	28		
	10	13	21			
CORSIM	1	11	19	15	29	
	2	12	16	15	19	
	3	17	18			
	4	14	29			
	5	11	15	14	15	16
	6	13	19	12	15	30
	7	12	15	15	30	
	8	12	18	15	30	
	9	12	17	34		
	10	13	21			

phase 4, and intersection 9 phase 3. Similar discrepancies can be observed in the results from each of the three TRANSYT-7F candidate models, although these results need not be listed at this time. Competing link effects are likely causing these discrepancies.

Because CORSIM is stochastic, the simulated queue lengths will vary from cycle to cycle. The possibility exists that during some cycles, traffic on a non-critical link will be heavier and extend the phase. This possibility becomes a high probability when one of the non-critical link flow ratios is almost as high as the critical link flow ratio, or when there are a high number of non-critical links present. On the other hand, when a non-critical link is not present, or when all non-critical links have very low flow ratios, this stochastic bias is minimized. This is reflected in the table 3-4 results because some of the

phases with multiple actuated links display no stochastic bias, whereas some of these phases are definitely affected.

The TRANSYT-7F candidate models are deterministic instead of stochastic. They are designed to analyze the critical link; however, they must be able to compensate for this competing link bias to produce more realistic results. One way to compensate is to use the queue calibration factor (f_q), documented within the Highway Capacity Manual appendix and in chapter 2, to account for randomness. The HCM [Transportation Research Board, 1997] states: "The queue calibration factor, f_q, was described by Akcelik as a factor required to account for randomness in arrivals in determining the average queue service time." The queue calibration factor equation, initially introduced in chapter 2 as equation (2-2), is repeated below:

$$f_q = 1.08 - 0.1\left(\frac{\text{actual green}}{\text{maximum green}}\right) \tag{2-2}$$

The experimental version of TRANSYT-7F was redesigned so as to apply the queue calibration factor to phases with multiple actuated links. This would not affect the results from experiment #1, in which the candidate models were showing good correlation with CORSIM, because in that experiment only one link was moving in each phase. Figure 3-15 contains partial output from the experimental version of TRANSYT-7F, showing some results generated by the first prototype model. This portion of the output shows how the queue calibration factor is applied under more complex conditions.

Links 706 and 708 in the first two phases are major-street left-turns. Since these links are the only links that can extend their own phase, no competing link bias exists,

Critlink = TRANSYT-7F critical link number
Start = starting time for this phase
MnG = minimum green time
Fq = queue calibration factor
G s= queue service time
Ge = green extension time
Y+R = yellow plus all red clearance time
Split = estimated phase time
Max = maximum phase time
Min = minimum phase time
Offset = offset

Critlink	Start	MnG	Fq	Gs	Ge	Y+R	Split	Max	Min	Offset
708	44	5	1	7.08	3.39	2	12	22	7	44
706	44	1	1	13.0	3.00	2	6	3	3	
0	59	5	1	0	0	2	49	76	30	
702	11	5	.98	12.1	3	2	15	15	7	
701	26	5	1.04	11.9	3.87	2	18	30	7	

Candidate Model Timing Plan: 12, 6, 49, 15, 18
Corsim Timing Plan: 12, 4, 50, 15, 19

Figure 3-15: Sample Numerical Effects of the Queue Calibration Factor (f_q)

and the queue calibration factor has no effect (f_q = 1.0). The third phase is the non-actuated major street phase on which calculating queue service time and green extension time is not necessary. The fourth phase is the minor street left-turn phase. It happens to be oversaturated in this example, and thus the phase time is equal to the maximum phase time. The fifth phase is a minor street through movement phase containing multiple actuated links, and its queue service time is multiplied or factored up by the queue calibration factor, 1.04.

In this example the final phase time reached 18 seconds after being increased by the queue calibration factor, but was still slightly lower than the average phase time observed in CORSIM. In fact, the table 3-4 results, showing a couple of seriously underestimated phase times, were produced while taking the queue calibration factor into

account. This suggests that the queue calibration factor equation is not fully compensating for competing link bias.

A likely reason for this is that the existing queue calibration factor formula only takes the actual green time and the maximum green time into account. This equation tends to produce higher queue calibration factor values as the difference between actual green time and maximum green time becomes higher. In other words, when the average phase time is already close to the maximum, randomness is minimized, and when the average phase time is much lower than the maximum, this increases the stochastic bias.

Although this may be appropriate, what does not seem appropriate is that the existing equation does not take into account the competing, non-critical link volumes or flow ratios. It was stated a few paragraphs earlier that the possibility of competing link bias becomes a high probability when one of the non-critical link flow ratios is almost as high as the critical link flow ratio, or when there are a high number of non-critical links present.

It appears that the queue calibration factor equation should be updated so as to account for traffic volumes or flow ratios on the non-critical link. Another possible strategy (for minimizing competing link bias) would be the computation of green extension times based on flow rates from all of the competing links, instead of just the critical link.

Signal Timing Optimization

Traditional signal timing strategies involved optimizing the splits (green times). However, with the advent of traffic actuated control, splits were no longer directly specified inside the controller. Instead a maximum green time (or force-off), a minimum

green time, and a gap setting were now specified in place of the green split. The result is that the green split can no longer directly be optimized.

The candidate models presented in this chapter are actually designed to estimate the actuated green split that can no longer be optimized. Why is it important to estimate the average green splits when instead the focus should be on optimization? The reason is that optimization programs like TRANSYT-7F need to know the correct, average amount of green time available on the arterial street in order to generate a design that can achieve progression. When the correct amount of green time available to the arterial street is known, this allows for intelligent optimization of the phasing pattern, cycle length, and offsets, although TRANSYT-7F itself does not currently optimize phasing pattern.

So proper optimization depends on knowing the correct actuated phase times, but in fact the phase times are likewise affected by the same signal settings (phasing pattern, cycle length, offset) that can be optimized. The result is a chicken-and-egg situation between phase time calculation and signal setting optimization. The existing model for actuated control within TRANSYT-7F does not account for this circular dependency, because it produces the same phase time estimates regardless of the optimal signal settings. However, because the candidate models are appropriately responsive to network-wide queue lengths and flow profiles, it should now be possible to analyze this circular dependency.

Theoretically, if a continuous series of phase time calibration runs and signal setting optimization runs converges to produce one unique network-wide signal timing plan, a new level of accuracy and optimality should be possible. This iterate-until-convergence concept is similar to the candidate models' strategy for calculating phase

times alone. Because the length of any actuated phase depends on the length of all other actuated phases, individual phase times should be continuously recalculated in response to one another until all are in agreement. Convergence is clearly achieved when additional recalculations do not change the phase times whatsoever. Subsequently the same concept can be applied to the bigger picture. Because the actuated phase lengths and optimal signal settings depend on each other, they should be continuously recalculated and re-optimized until all are in agreement. Convergence is clearly achieved when additional recalculations do not change the phase times whatsoever, and additional re-optimizations do not improve network performance whatsoever. Figure 3-16 illustrates this phenomenon.

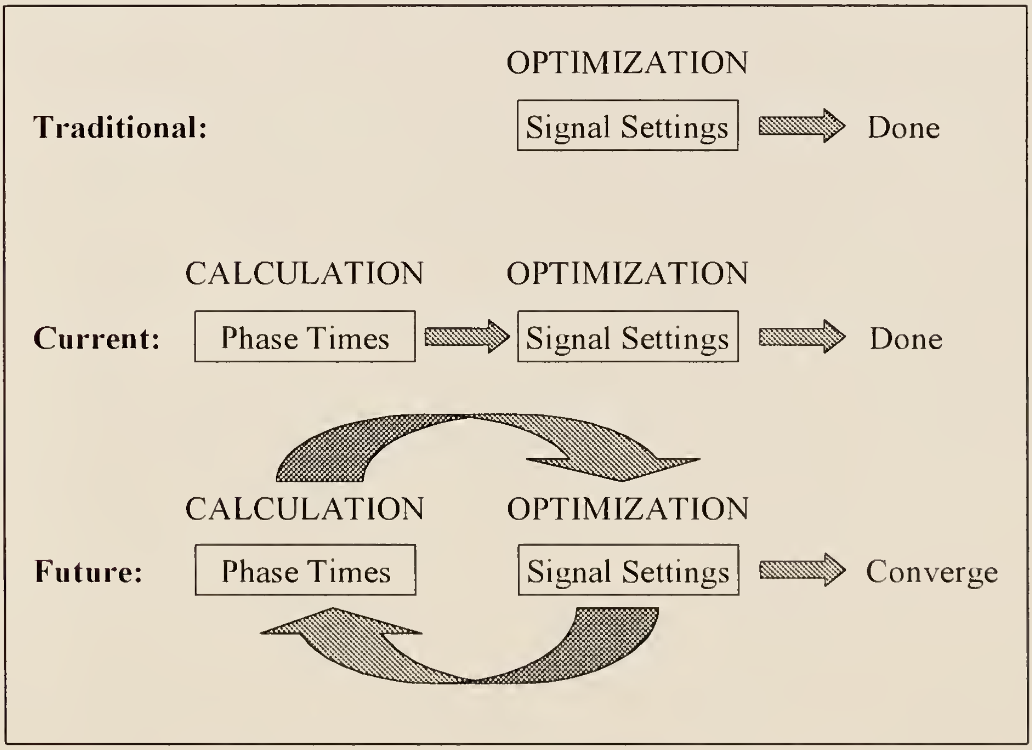


Figure 3-16: Evolution of Signal Timing Optimization Strategies

Chapter Three Summary

This chapter described the development of new models for improvement of TRANSYT-7F performance. It was stated that the queue service time — green extension time model structure from the literature is preferable to the target degree of saturation strategy. Advantages and disadvantages of the existing queue service time models were listed. Queue service time model disadvantages are elegantly eliminated by adopting the information embedded within TRANSYT-7F simulation. This chapter also described the development of new models for estimating green extension time that are not constrained to the assumption of random vehicle arrivals. Green extension times calculated by the first prototype model are responsive to maximum green and progression effects. The second prototype model provides an additional alternative to existing models (from the literature) that operate under the assumption of random arrivals.

Additional issues arise when applying these models to compute network-wide signal timing plans. The relationship between queue service time and green extension time is fundamentally different when comparing isolated and coordinated operations. Because of early return to green effects, allocation of unused green time must also be conducted in a different manner under coordinated operations. Overlap phasing effects undermine the traditional techniques for determining critical actuated links. Thus, it is necessary to perform the "critical link search", and continuously evaluate each available link that could possibly govern phase termination. Competing link effects introduce a bias into the results between the deterministic candidate models and the microscopic simulation models. The queue calibration factor (f_q), or the calculation of green

extension times based on flow rates from all competing links, are potential strategies for minimizing competing link bias. Finally, the new candidate models theoretically change the way signal timing optimization can be conducted. Theoretically, if actuated phase times are responsive to system effects, then multiple phase time calibrations and signal setting optimizations may result in network wide convergence, and a solution that is more likely to materialize when applied in the field.

At this time, all candidate models must be tested in more detail to further ascertain their effectiveness. Chapter 4 describes the results of several tests. These tests compare results of the TRANSYT-7F existing and candidate actuated control models against results from the CORSIM program.

CHAPTER 4
MODEL TESTING

In this chapter, models will be tested as candidates for improvement of TRANSYT-7F performance. Tasks 7 & 8 mentioned in chapter 1 will be carried out in this chapter. The testing plan is described at the beginning of the chapter, whereas testing results are described in the remainder of the chapter. Some of the models tested are found in the literature and were described in chapter 2. The other models are prototypes developed solely for the purpose of this study as described in chapter 3. These models meet the minimum requirements for an improved model as described in chapter 1, but must now be tested more thoroughly in order to gauge strengths and weaknesses, advantages and disadvantages, accuracy and practicality, etc.

The primary purpose of this chapter is to evaluate green extension time models. The results of this chapter allow for direct comparison between numerous models for green extension time, because the queue service time model is “held constant”, or unchanged within any one experiment. In chapters 2 and 3, the advantages of computing queue service time based on the queue accumulation polygon were discussed at length, so there is little question about how to correctly model queue service time.

When multiple intersections are involved, it becomes necessary to extract queue service time directly from the results of TRANSYT-7F simulation, in order to account for permitted left-turn, progression, and spillback effects.

List of Candidate Models

In chapter 4, seven models will be tested and compared against each other as well as CORSIM. It is necessary to choose specific, brief names for the models to be identified when presenting large quantities of testing results. Because small names are not descriptive enough to describe the individual models, acronyms will instead be used to describe the models. Table 4-1 presents the list of candidate model names, and also presents an overview of the candidate model features.

Table 4-1: Candidate Model Naming Convention and Characteristics

Name	GP1	GP2	GP3	GL1	GL2	DL1	DL2
Applies queue service -- green extension strategy	*	*	*	*	*		
Applies target degree of saturation strategy						*	*
Prototype model developed during this study	*	*	*				
Existing model within the literature				*	*	*	*
Developed under the assumption of random arrivals				*	*		*
Developed under the assumption of uniform arrivals		*	*				
Developed to reflect the flow profile	*						
actuated signal settings	*	*	*	*	*		*
detector configuration				*	*		*
Applies a methodology of iteration and convergence	*	*		*	*		*
Applies a uniform queue service time model			*				
Model developer	A	A	A	B	C	D	B
A: Hale							
B: Akcelik							
C: Husch							
D: Skabardonis							

Testing Strategy

In chapter 2 it was stated that the CORSIM program would be useful as a benchmark for comparisons between traffic-actuated control models. CORSIM is used extensively throughout this chapter for that very purpose. Although CORSIM is not perfect and identical to the real world, it is understood that, if TRANSYT-7F can compute actuated phase times that show good correlation with those from CORSIM, this would indicate a substantial improvement over the existing model.

In order to test the candidate models more thoroughly, an experimental version of TRANSYT-7F was created in order to implement them. In chapter 3, it was stated that additional program development was required in order to achieve proper model implementation. To properly account for early return to green effects, a new green time reallocation algorithm was implemented. To account for overlap phasing effects, the critical link search technique was implemented. To account for competing link effects, application of the queue calibration factor was modified. It was applied selectively, and only on phases that are vulnerable to competing link bias.

Another point to note is that all test conditions and data from this chapter are purely hypothetical and fictitious. Test conditions were not based on those actually observed in the field. Rather a wide variety of artificial data was created and processed so as to produce a comprehensive set of useful results. Although the artificial data may not be identical to that observed at any one real world location, the amount and range of data compiled allow the results to reflect various conditions from the field. The testing data included a variety of:

- traffic volumes and degrees of saturation
- phasing patterns and signal timing plans
- lane configurations and channelizations
- offsets, progression, and flow profiles
- left-turn protection (protected-only, permitted-only, compound protection)
- vehicle turning movement percentages

To gain perspective, it is also important to be aware of the limitations of testing that was conducted. Indeed, it can be difficult to confidently evaluate testing results without being aware of these limitations. In the chapter 4 testing, the following conditions were not analyzed:

- shared lanes served by permitted, actuated phases
- a grid, or a network of signals
- phase skipping
- detector setback
- intersections with 5 or more approaches
- volume-density controllers, or controllers with gap reduction functionality
- wide varieties of internal link lengths
- multi-cycle or multi-period simulation

Chapter 5 discussions will suggest that some of these testing limitations listed above are inconsequential, whereas some of these limitations deserve to be scrutinized by future research. Within the specified set of limitations, the results of chapter 4 testing should prove useful.

Calibration of CORSIM and TRANSYT-7F

In order to test the candidate models, each one was programmed so as to operate in conjunction with TRANSYT-7F. This means that in order to execute one of the candidate models, the experimental program was used to process typical TRANSYT-7F input files. Assuming that the candidate models were programmed correctly, the biggest remaining obstacle to compiling useful data and information becomes the differences between TRANSYT-7F and CORSIM themselves.

For example, one possible pitfall would be to observe large differences between results from a candidate model and CORSIM, and conclude the candidate model has poor performance, not realizing that the differences were actually caused by inappropriate differences in the input data for both programs. Keeping the two programs consistent with each other becomes more difficult when they use different input variables to model the same processes. A few of these input parameters affecting phase times include:

1. Queue Discharge Headway (CORSIM), Saturation Flow Rate (TRANSYT-7F)
2. Start-up Lost Time
3. Amber Response (CORSIM), Extension of Effective Green (TRANSYT-7F)
4. Internal & External Link Length
5. Free Flow Speed
6. Lane Usage, Channelization, or Designation
7. Entry Node and Approach Volumes (CORSIM), Link Volumes (TRANSYT-7F)

In the many experiments described in this chapter, and throughout the data compilation effort, careful attention was paid to ensure that biases between these two major programs would be kept to a minimum. The details of this calibration effort are discussed in the appendix.

Chapter Four List of Experiments

This chapter will describe numerous experiments designed to scrutinize various aspects of the candidate models. The initial experiments involve simplistic conditions in order to confirm the basic validity of the models, to begin learning about basic differences in the results that they produce, and to shed some light on basic differences between coordinated and fully actuated control. Subsequent experiments introduce additional complications one at a time, in order to scrutinize specific aspects of the candidate models.

1a. Single Intersection & Background Cycle Length (Semi-Actuated)
1b. Single Intersection & No Background Cycle Length (Fully-Actuated)
1 supplemental. Semi-Actuated Queue Service Time Feedback Effect
2a. Arterial Progression Effects (Internal Links)
2b. Arterial External Links
2 supplemental. Arterial Optimization Effects
3a. Arterial Permitted Left-turn Effects (with CORSIM Defaults)
3b. Arterial Permitted Left-turn Effects (with CORSIM Calibration)
4. Optimization and Spillback Effects

Single Intersection Testing

The first model testing experiment involved the simplest possible conditions, namely a single signalized intersection having only two phases. When a typical 4-leg signalized intersection has only two phases, this means that a “green ball” is first displayed in one direction (e.g. north-south) during the first phase, and subsequently displayed in the other direction (e.g. east-west) during the second phase. The idea behind this simplistic experiment is that errors in programming, or obvious differences between

the candidate models, could be eliminated or identified early on before conducting tests with more complicated conditions.

The first experiment was actually divided into two parts. In the first part, the minor street phase is actuated, the major street phase is non-actuated, and the background cycle length is assumed to be 80 seconds. In the second part of the experiment, both phases are actuated, and so the cycle length is affected by the individual phase times. A diagram of the test intersection is illustrated in figure 4-1.

Experiment #1a (Background Cycle Length)

In this experiment sensitivity analysis is performed to gauge the candidate models' response to variations in the gap setting and traffic volume only. In subsequent, more complicated experiments, the models' response to variations in other conditions will be tested as well. Although figure 4-1 illustrates identical lane configurations on each approach, traffic flow was only defined on two of the four approaches (northbound and eastbound) to keep the experiment as simple as possible. The simplistic phasing diagram is illustrated in figure 4-2, and the east-west direction moving left to right across the page is assumed to be the major arterial street.

In this first part of experiment #1, since the background cycle length is assumed to be 80 seconds, this means that the length of the non-actuated phase will always be equal to 80 minus the length of the actuated phase. Although the length of the non-actuated phase will affect results, the main purpose here is to scrutinize the candidate models' prediction of the actuated phase lengths, presented in table 4-2.

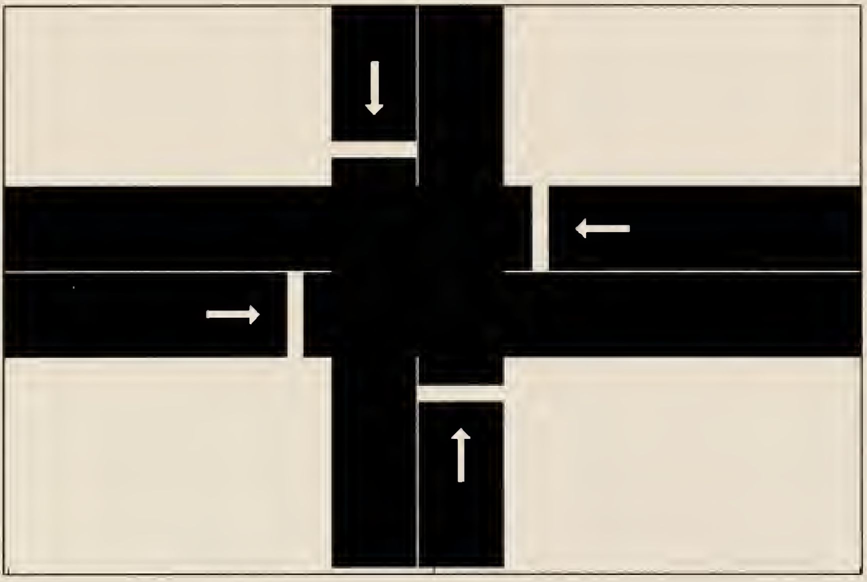


Figure 4-1: Experiment #1 Intersection Configuration

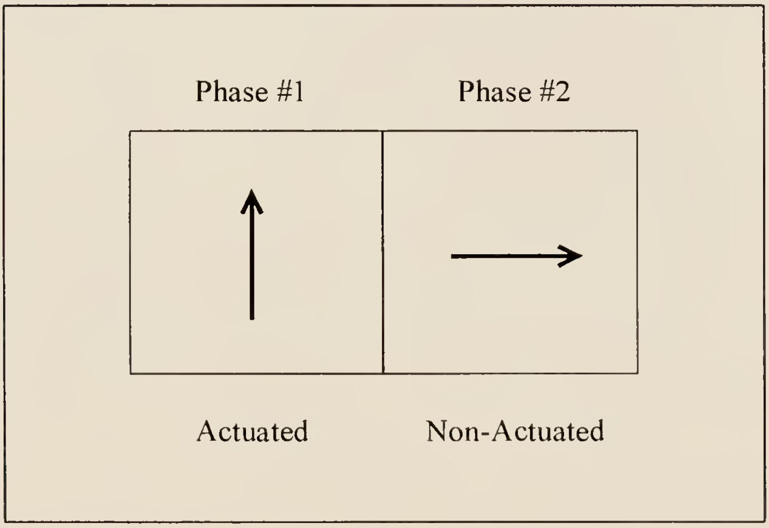


Figure 4-2: Experiment #1a Phasing Diagram

Table 4-2: Experiment #1a Actuated Phase Lengths

				Coordination in Effect Actuated Phase Lengths				
300vph Gap	DL1	CORSIM	GL1	GP2	GL2	GP1	DL2	GP3
2	18.0	19.6	20.3	18.8	18.6	19.0	25.0	19.5
3	18.0	20.7	21.5	19.8	19.5	20.4	25.0	20.9
4	18.0	21.9	22.6	20.9	20.4	21.8	26.0	22.2
5	18.0	23.4	24.1	22.0	21.2	23.4	26.5	23.5
500vph Gap	DL1	CORSIM	GL1	GP2	GL2	GP1	DL2	GP3
2	28.0	27.9	28.6	27.4	27.1	27.6	32.0	28.5
3	28.0	29.4	30.1	28.3	27.7	29.2	33.0	29.8
4	28.0	31.1	31.6	29.4	28.5	31.3	34.0	31.3
5	28.0	33.0	33.3	30.5	30.0	33.3	34.5	32.9
700vph Gap	DL1	CORSIM	GL1	GP2	GL2	GP1	DL2	GP3
2	39.0	36.7	37.5	35.6	35.9	36.3	38.0	37.5
3	39.0	37.5	38.7	36.5	36.2	38.3	39.0	39.0
4	39.0	39.4	40.6	37.6	37.6	40.4	40.0	40.7
5	39.0	42.5	43.0	38.7	38.7	43.0	41.0	42.5

The existing actuated control model within TRANSYT-7F is deterministic, as are the candidate models, and will thus produce the same average phase time estimate every time. CORSIM, on the other hand, is stochastic, and so its average phase time results are affected by the number of simulated cycles. In general, a higher number of simulated cycles provides an average phase time value that is more indicative of CORSIM's behavior [Hale, 1997]. In the first part of experiment #1, 3600 seconds (45 cycles) were simulated in CORSIM for each data point, and the average phase time was computed to tenths of a second precision.

In table 4-2, the gap setting does not affect phase time estimates from the TRANSYT-7F release 8 existing model (DL1). This is not appropriate since increases in

the gap setting should always increase the effective lost time within the cycle, and thus increase the phase time. In order to interpret overall results more clearly it is helpful to observe the percentage of error (relative to CORSIM) for each data point in table 4-3.

Table 4-3: Experiment #1a Percent Error (Relative to CORSIM)

300vph Gap	DL1	GL1	GP2	GL2	GP1	DL2	GP3
2	-8.2	3.6	-4.1	-5.1	-3.1	27.6	-0.6
3	-13.0	3.9	-4.3	-5.8	-1.4	20.8	1.0
4	-17.8	3.2	-4.6	-6.8	-0.5	18.7	1.4
5	-23.1	3.0	-6.0	-9.4	0.0	13.2	0.4
500vph Gap	DL1	GL1	GP2	GL2	GP1	DL2	GP3
2	0.4	2.5	-1.8	-2.9	-1.1	14.7	2.2
3	-4.8	2.4	-3.7	-5.8	-0.7	12.2	1.4
4	-10.0	1.6	-5.5	-8.4	0.6	9.3	0.6
5	-15.2	0.9	-7.6	-9.1	0.9	4.5	-0.3
700vph Gap	DL1	GL1	GP2	GL2	GP1	DL2	GP3
2	6.3	2.2	-3.0	-2.2	-1.1	3.5	2.2
3	4.0	3.2	-2.7	-3.5	2.1	4.0	4.0
4	-1.0	3.0	-4.6	-4.6	2.5	1.5	3.3
5	-8.2	1.2	-8.9	-8.9	1.2	-3.5	0.0

Based on the results from table 4-3, candidate models that apply the queue service time — green extension time strategy appear to calculate actuated phase times having better correlation with those from CORSIM, relative to candidate models that apply the target degree of saturation strategy. In addition to compiling average phase times during

experiment #1, additional data was compiled on queue service times (g_s) and green extension times (g_e) for the respective models. These results are listed in table 4-4. Note that in the results of experiment #1, queue service time is assumed to include start-up lost time, whereas the Highway Capacity Manual keeps these two parameters separate in its calculations.

Table 4-4: Experiment #1a Queue Service and Green Extension Times

300vph												
Gap	CORSIM		GL1			GP2		GL2		GP1		DL2
	gs	ge	gs	ge	gs	gs	ge	gs	ge	gs	ge	gs
2	15.0	2.6	14.4	3.9	14.6	15.3	2.2	14.6	2.0	14.6	2.4	23.0
3	15.1	3.7	14.2	5.3	14.4	15.3	3.4	14.5	3.0	14.4	4.0	23.0
4	14.9	5.0	13.8	6.8	14.2	15.3	4.7	14.4	4.0	14.0	5.8	24.0
5	14.9	6.6	13.6	8.5	14.0	15.3	6.0	14.2	5.0	13.8	7.6	24.5
500vph												
Gap	CORSIM		GL1			GP2		GL2		GP1		DL2
	gs	ge	gs	ge	gs	gs	ge	gs	ge	gs	ge	gs
2	23.2	2.7	22.3	4.3	23.1	24.2	2.3	23.1	2.0	22.9	2.7	30.0
3	22.2	5.2	22.0	6.1	22.7	24.2	3.6	22.7	3.0	22.3	4.9	31.0
4	21.4	7.7	21.4	8.2	22.3	24.2	5.1	22.5	4.0	21.6	7.7	32.0
5	19.6	11.4	20.6	10.7	21.8	24.2	6.7	22.0	6.0	20.8	10.5	32.5
700vph												
Gap	CORSIM		GL1			GP2		GL2		GP1		DL2
	gs	ge	gs	ge	gs	gs	ge	gs	ge	gs	ge	gs
2	30.3	4.4	30.6	4.9	31.2	33.1	2.4	31.9	2.0	31.2	3.1	36.0
3	28.9	6.6	29.3	7.4	30.6	33.1	3.9	31.2	3.0	30.0	6.3	37.0
4	27.9	9.5	28.3	10.3	30.0	33.1	5.6	30.6	5.0	28.3	10.1	38.0
5	26.4	14.1	26.8	14.2	29.3	33.1	7.4	29.7	7.0	26.8	14.2	39.0

The GP1 and GL1 models produce green extension times having good correlation with those from CORSIM. Green extension times from other models have weaker correlation. In addition, queue service time and green extension time appear to be inversely related. That is, when queue service time increases, green extension time decreases, and vice versa. These results will be discussed further after results from the second part of experiment #1 are presented.

Experiment #1b (Fully Actuated)

The second part of this experiment also uses the intersection geometry from figure 4-1. Once again traffic flow was only defined on two of the four approaches (northbound and eastbound) to keep the experiment as simple as possible. Since both phases are actuated in this part of the experiment, notation on the simplistic phasing diagram is updated and illustrated in figure 4-3.

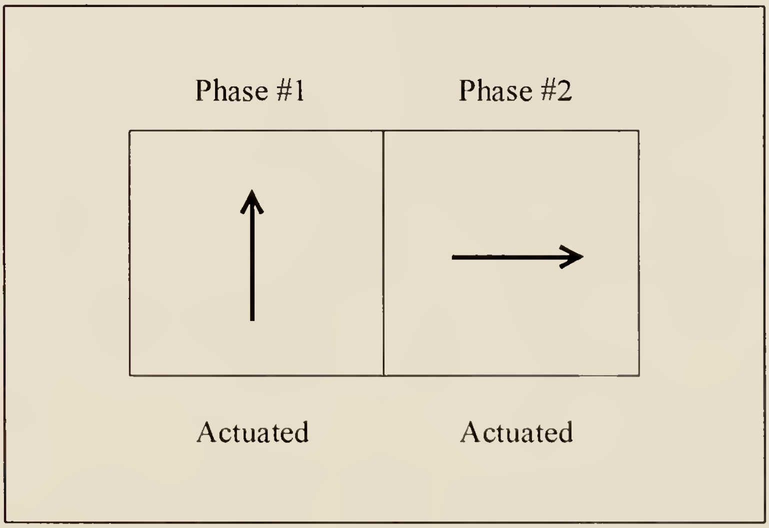


Figure 4-3: Experiment #1b Phasing Diagram

In this second part of experiment #1, the cycle length is affected by the individual actuated phase lengths. Traffic volume and gap setting variations are analyzed as they were in the previous experiment. It is important to note that input conditions on both approaches are identical, meaning that each actuated phase length will equal 1/2 of the cycle length results presented in table 4-5. In this table, no results are available for DL1, because this model was not designed for fully actuated control. Once again 3600 seconds were simulated in CORSIM to obtain each data point, although the number of simulated cycles varied due to average cycle length variations.

Table 4-5: Experiment #1b Actuated Cycle Lengths

Isolated Conditions Actuated Cycle Lengths							
300vph Gap	CORSIM	GL1	GP2	GL2	GP1	DL2	GP3
2	17.4	20.5	16.3	16.0	17.1	12.0	19.0
3	20.5	24.1	19.5	18.4	21.2	12.0	22.1
4	23.1	28.0	22.5	21.2	25.6	12.0	26.0
5	27.7	32.1	26.1	23.6	29.9	12.0	30.1
500vph Gap	CORSIM	GL1	GP2	GL2	GP1	DL2	GP3
2	25.8	29.7	23.3	22.0	24.1	31.0	29.1
3	30.2	35.6	27.5	25.5	31.7	33.0	35.2
4	38.1	42.1	32.1	29.1	41.0	35.0	41.4
5	44.5	50.0	37.5	35.4	49.8	37.0	49.1
700vph Gap	CORSIM	GL1	GP2	GL2	GP1	DL2	GP3
2	49.8	57.3	43.2	41.2	47.2	50.0	59.4
3	59.2	70.9	51.3	47.0	64.9	54.0	73.1
4	79.2	87.0	61.1	57.4	87.2	58.0	89.1
5	92.7	107.6	71.1	69.0	111.0	62.0	105.3

In order to interpret these results more clearly it is helpful to observe the percentage of error (relative to CORSIM) for each data point in table 4-6.

Table 4-6: Experiment #1b Percent Error (Relative to CORSIM)

300vph Gap	GL1	GP2	GL2	GP1	DL2	GP3
2	18	-6	-8	-2	-31	9
3	18	-5	-10	3	-42	8
4	21	-3	-8	11	-48	13
5	16	-6	-15	8	-57	9
500vph Gap	**GL1**	**GP2**	**GL2**	**GP1**	**DL2**	**GP3**
2	15	-10	-15	-7	20	13
3	18	-9	-16	5	9	17
4	11	-16	-24	8	-8	9
5	12	-16	-20	12	-17	10
700vph Gap	**GL1**	**GP2**	**GL2**	**GP1**	**DL2**	**GP3**
2	15	-13	-17	-5	0	19
3	20	-13	-21	10	-9	23
4	10	-23	-27	10	-27	13
5	16	-23	-26	20	-37	14

In tables 4-5 and 4-6, candidate model correlation with CORSIM appears to get worse when vehicle inter-arrival times (5.14 seconds for 700 vph) approach the gap setting (e.g. 5 seconds). Additional data compiled on queue service times (g_s) and green extension times (g_e) for the respective models are listed in table 4-7. In this table, queue

service time and green extension time are positively related. That is, when green extension time increases, queue service time also increases.

Table 4-7: Experiment #1b Queue Service and Green Extension Times

300vph												
Gap	CORSIM		GL1			GP2		GL2		GP1		DL2
	gs	ge	gs	ge	gs	gs	ge	gs	ge	gs	ge	gs
2	3.5	3.2	4.4	3.9	4.0	5.3	2.2	4.0	2.0	4.2	2.4	4.0
3	4.1	4.2	4.8	5.3	4.4	5.7	3.4	4.2	3.0	4.6	4.0	4.0
4	4.2	5.4	5.2	6.8	4.6	6.3	4.7	4.6	4.0	5.0	5.8	4.0
5	4.9	7.0	5.6	8.5	5.0	7.0	6.0	4.8	5.0	5.4	7.6	4.0
500vph												
Gap	CORSIM		GL1			GP2		GL2		GP1		DL2
	gs	ge	gs	ge	gs	gs	ge	gs	ge	gs	ge	gs
2	8.1	2.8	8.5	4.3	7.4	10.3	2.3	7.0	2.0	7.4	2.7	13.5
3	8.5	4.6	9.7	6.1	8.2	12.0	3.6	7.8	3.0	8.9	4.9	14.5
4	10.2	6.9	10.8	8.2	8.9	13.6	5.1	8.5	4.0	10.8	7.7	15.5
5	11.6	8.7	12.3	10.7	10.0	15.8	6.7	9.7	6.0	12.3	10.6	16.5
700vph												
Gap	CORSIM		GL1			GP2		GL2		GP1		DL2
	gs	ge	gs	ge	gs	gs	ge	gs	ge	gs	ge	gs
2	18.7	4.2	21.7	4.9	17.2	25.3	2.4	16.6	2.0	18.5	3.1	23.0
3	21.5	6.1	26.1	7.4	19.8	30.7	3.9	18.5	3.0	24.2	6.3	25.0
4	27.4	10.2	31.2	10.3	23.0	37.0	5.6	21.7	5.0	31.2	10.4	27.0
5	30.0	14.4	37.6	14.2	26.1	43.2	7.4	25.5	7.0	38.9	14.6	29.0

Experiment #1 Analysis

At this time conclusions will be drawn (if possible) from the data that was compiled during experiments #1a and #1b.

Target degree of saturation models

The first part of this experiment showed that the existing model within TRANSYT-7F (DL1) does not respond appropriately to changes in the gap setting (observation #1). The second part of experiment #1 was simply a reminder that the existing model cannot be used to analyze isolated intersections. Although TRANSYT-7F was not primarily designed to focus on isolated intersections, the overall results from experiment #1 indicate that the existing model leaves plenty of room for improvement.

Experiment #1 also indicated poor performance from Akcelik's iterative target degree of saturation model (DL2). Although the performance of this model was expected to decline in the presence of complex network conditions, since no methodology exists for adjusting the target degree of saturation in response to network-wide traffic flow effects, the poor performance of this model under simple conditions was unexpected.

Estimating phase times under high volume and high gap settings

Correlation was poor between all models when vehicle inter-arrival times approached the gap setting. This result is not as troublesome as it may appear, for several reasons. Without a relatively low maximum green in effect, this becomes an unstable mode of operation where a very high number of vehicle extensions can occur, and the smallest biases between the models are likely to become magnified. Chapter 3 states that the queue service time structure used by the candidate models is proven and robust, so minor biases between CORSIM and TRANSYT-7F simulation could be getting magnified. Also, rarely in the real world do vehicle inter-arrival times approach the gap setting. With a typical gap setting of 3 seconds, an unusually high volume near 1200 vph per lane would be necessary for conditions to become unstable. Under these conditions,

there would probably be a reasonable maximum green time in place to keep the operation under control.

Candidate model tendencies

The NCHRP green extension time model (GL1), also developed by Akcelik, appears to estimate high green extension times and phase times, relative to the other models. Conversely, the prototype green extension time model based on uniform vehicle arrivals (GP2) appears to estimate low green extension times and phase times, relative to the other models.

This is consistent with the nature of these models as described in chapters 2 and 3. In the HCM [Transportation Research Board, 1997], the NCHRP green extension time model is documented as recognizing the bunched negative exponential distribution. This distribution is similar to the regular negative exponential distribution, associated with random or Poisson vehicle arrivals. Random arrivals would tend to extend a phase more often than truly uniform arrivals, under which only one vehicle extension is mathematically possible. Therefore, it is not surprising that the NCHRP model would predict relatively high phase times, and the prototype model based on uniform arrivals would predict relatively low phase times.

Joint Poisson probability model

Husch's joint Poisson probability model (GL2, for estimating green extension time) consistently undershoots even the prototype model based on uniform arrivals, whose phase times are already lower than those from CORSIM or any other candidate model. This is consistent with observations from chapters 2 and 3, where Husch's model

was shown to assume immediate phase termination even when only a 50% chance of phase termination exists.

Candidate model performance

Prototype green extension time models GP1 and GP3 had the best correlation with CORSIM under simple coordinated conditions. It should be noted that the NCHRP GL1 model from Akcelik had good correlation under these conditions also. Coordinated conditions are more important than isolated conditions in the scope of network analysis using programs like CORSIM and TRANSYT-7F. Prototype model GP1, which is based on applied probability analysis of the flow profile, also showed the best correlation under simple isolated conditions. This result was unexpected since the other candidate models were specifically developed with random and uniform arrivals in mind. Random or uniform vehicle arrivals are expected at isolated intersections. Thus the GP1 model exhibits the best overall correlation with CORSIM so far; however, it is desirable to test the candidate models under more complex conditions.

Comparison of coordinated vs. isolated actuated control behavior

The experiment #1 results indicate an interesting difference in behavior based on whether or not a background cycle length is in effect. When there is a specific cycle length being enforced, queue service time and green extension time are inversely related. When there is no cycle length being enforced, queue service time and green extension time are positively related.

Certainly the gap setting has the ability to affect the phase time as well as the green extension time under coordinated conditions. The results from tables 4-2 and 4-4 demonstrate that phase times increase noticeably and consistently along with green

extension times when the gap setting is increased. A minor amount of queue service time feedback and reduction does occur (as evidenced in the table 4-3 results), but not enough to prevent definite increases in the phase time. This may be happening because the gap setting affects the total lost time, i.e., time during which vehicles cannot be served during the cycle, whereas the other parameters with the ability to increase green extension time do not affect the total lost time.

In order to shed some light on whether certain other parameters (e.g., detector length, approach speed, vehicle length, maximum green) have any effect on phase times under coordinated conditions, a quick test was performed to supplement the results of experiment #1. Tables 4-2 and 4-4 showed that the average phase time was 29.4 seconds given an hourly volume of 500 vehicles and a gap setting of 3 seconds, and this data point was chosen to test whether the phase time would change noticeably when varying the other parameter values in CORSIM. Table 4-8 shows that, although the queue service time — green extension time feedback effect prevents major shifts in the phase time, there are minor shifts. Based on this it appears that actuated phase times under coordination cannot be assumed to be "unmovable".

Table 4-8: Experiment #1 Supplemental – Phase Time Shifting under Coordination

Max Green (seconds)	Det Length (feet)	Speed (mph)	Veh Length (feet)	Gs (seconds)	Ge (seconds)	Y+R (seconds)	Phase Time (seconds)
90	**30**	**25**	**16**	**22.2**	**5.2**	**2**	**29.4**
28				22.3	3.1		27.4
	90			21.2	6.7		29.9
		45		19.8	6.3		28.1
			20	19.9	7.3		29.2

Experiment #1 Conclusions

In the upcoming experiments, testing will be conducted using more complex analysis conditions. To summarize, here are some reasonable conclusions from the first experiment:

1. Candidate models that apply the queue service time — green extension time strategy demonstrate superior performance, relative to the candidate models that apply the target degree of saturation strategy, given simple conditions.
2. The behavior and numerical relationship between queue service time and green extension time is affected by whether or not a background cycle length is imposed.

Arterial Street Testing

The candidate models were also tested on their ability to analyze an arterial street with multiple signalized intersections. Arterial analysis provides insight into factors that can have a significant impact on traffic operations, including progression, non-uniform vehicle arrivals, and spillback. Once again the hope is that candidate TRANSYT-7F models will achieve better correlation with CORSIM, a program that is recognized for its accurate simulation but performs no signal timing optimization. Once its simulation of actuated control shows better agreement with CORSIM, TRANSYT-7F's optimization results should benefit greatly.

Data sets and input files for CORSIM and TRANSYT-7F in the upcoming experiments are based those developed for a University of Florida project for developing traffic engineering benchmark data [Henry, 1999]. The arterial street conditions from the

benchmark data sets contain differing operating characteristics at each individual intersection, including traffic volume, degree of saturation, lane channelization, phase sequence, and left-turn treatment. This will be useful in learning about how the models perform under more complex conditions. Figure 4-4 illustrates this hypothetical arterial street that was used for testing purposes.

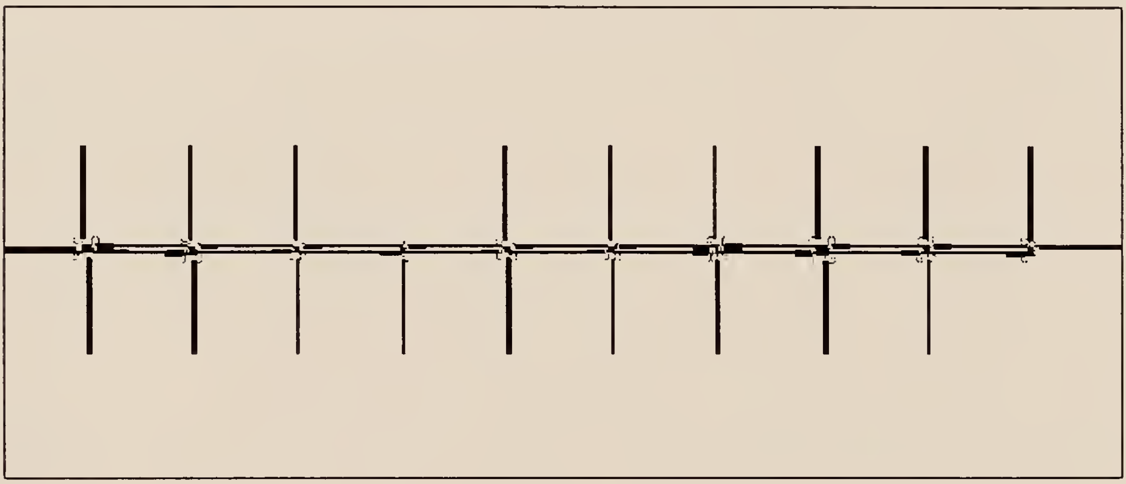


Figure 4-4: Geometry of the Arterial Street Used throughout Chapter 4 Testing

Experiment #2, Progression Effects

An important part of arterial street testing involves taking a look at if or how the pattern of vehicle arrivals affects traffic-actuated operations. When performing experiments with a single intersection only, the effects of metered arrivals or progressed flow were ignored. Progression effects would be caused by the presence of a nearby upstream traffic signal. Instead, it was possible to assume that vehicle arrivals were uniform or random in nature. Figure 4-5 illustrates the links that are susceptible to progression effects.

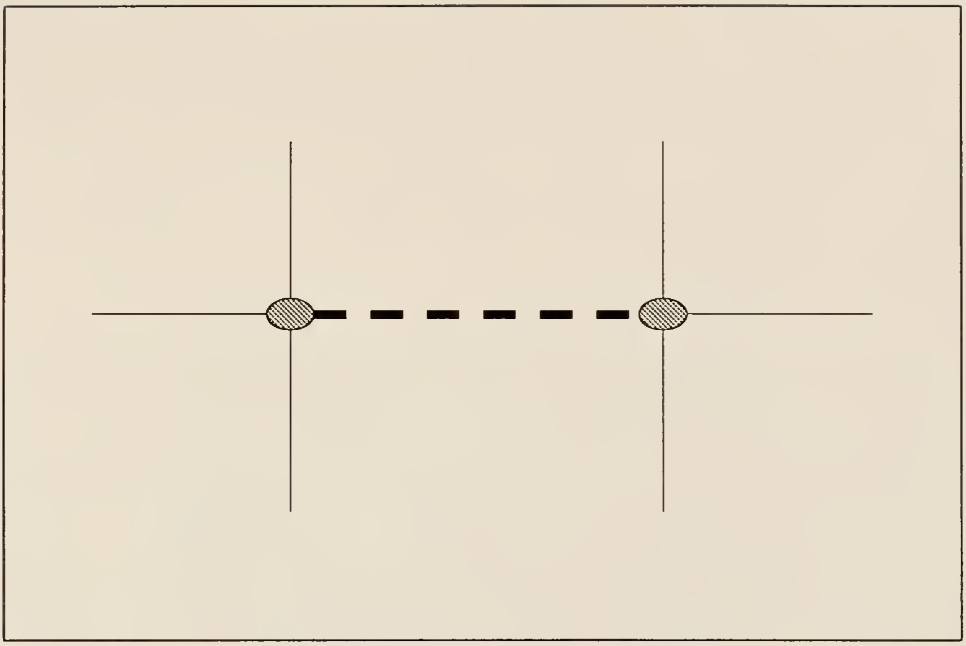


Figure 4-5: Internal Links Susceptible to Progression Effects

The dashed lines indicate “internal” links whose phase times may be susceptible to non-uniform vehicle arrivals or progression effects. In the figure this dashed line looks like perhaps only one link. However, by TRANSYT-7F terminology, the single dashed line contains multiple links because each turning movement is represented by a unique link. Thus, if all three turning movements are being made at both intersections in figure 4-5, there are at least six links here, possibly more under complex conditions. By CORSIM terminology, the single dashed line contains two uni-directional links, e.g. (1-2) and (2-1) if the nodes were numbered 1 and 2. The regular lines indicate “external” links where vehicle arrivals are typically assumed to be uniform or random.

External links can be thought of as coming out of a shopping mall, or residential area, or any area with no nearby upstream signal. This means that vehicle arrivals may also be assumed uniform or random, and links assumed external, when the distance

between traffic signals becomes sufficiently large such that platoons of vehicles have enough time and space to spread out. However, it is unclear exactly what distance is sufficient to dissipate the platoons and render progression effects negligible. This distance is hypothesized to be somewhere between 1 and 3 miles.

In the arterial street benchmark data set, the distance between signals is 1000 feet, and so progression effects are expected on the internal links. Because the major street non-actuated phase is almost always present on internal links, this leaves only the actuated major street left-turn phases as ones that could possibly be influenced by progression effects. Although figure 4-5 illustrates only two intersections that would have a maximum of four internal left-turn phases to analyze, the benchmark data set contains 10 intersections and 14 internal left-turn phases.

Testing conditions: experiment #2 for evaluating progression effects

To set the stage for experiment #2, a second arterial street data set was generated with a new offset design in order to obtain more data and more information on progression. Although the original data set contained offsets directly from the benchmark data study, the second data set contained offsets generated from an offsets-only optimization run in TRANSYT-7F release 8. These two sets of offsets were subsequently translated into yield points for use within CORSIM. Table 4-9 shows that the second offset design is significantly different and should be effective for creating unique traffic flow profiles and additional data. Notice that the offset differences between neighboring intersections are not always equal to the yield point differences. This discrepancy occurs because in the TRANSYT-7F data sets, although offsets may be

referenced to any phase, in this case they are referenced to the end of an actuated phase. CORSIM yield points are always referenced to the end of the non-actuated phase.

Table 4-9: Multiple Offset Designs Used to Test Progression Effects

Original Offsets	37	56	16	35	55	14	33	53	12	32
New Offsets	81	14	29	28	11	12	48	59	4	44
Original Yield Points	90	9	74	3	8	67	86	6	70	0
New Yield Points	34	67	87	96	64	65	1	12	62	12

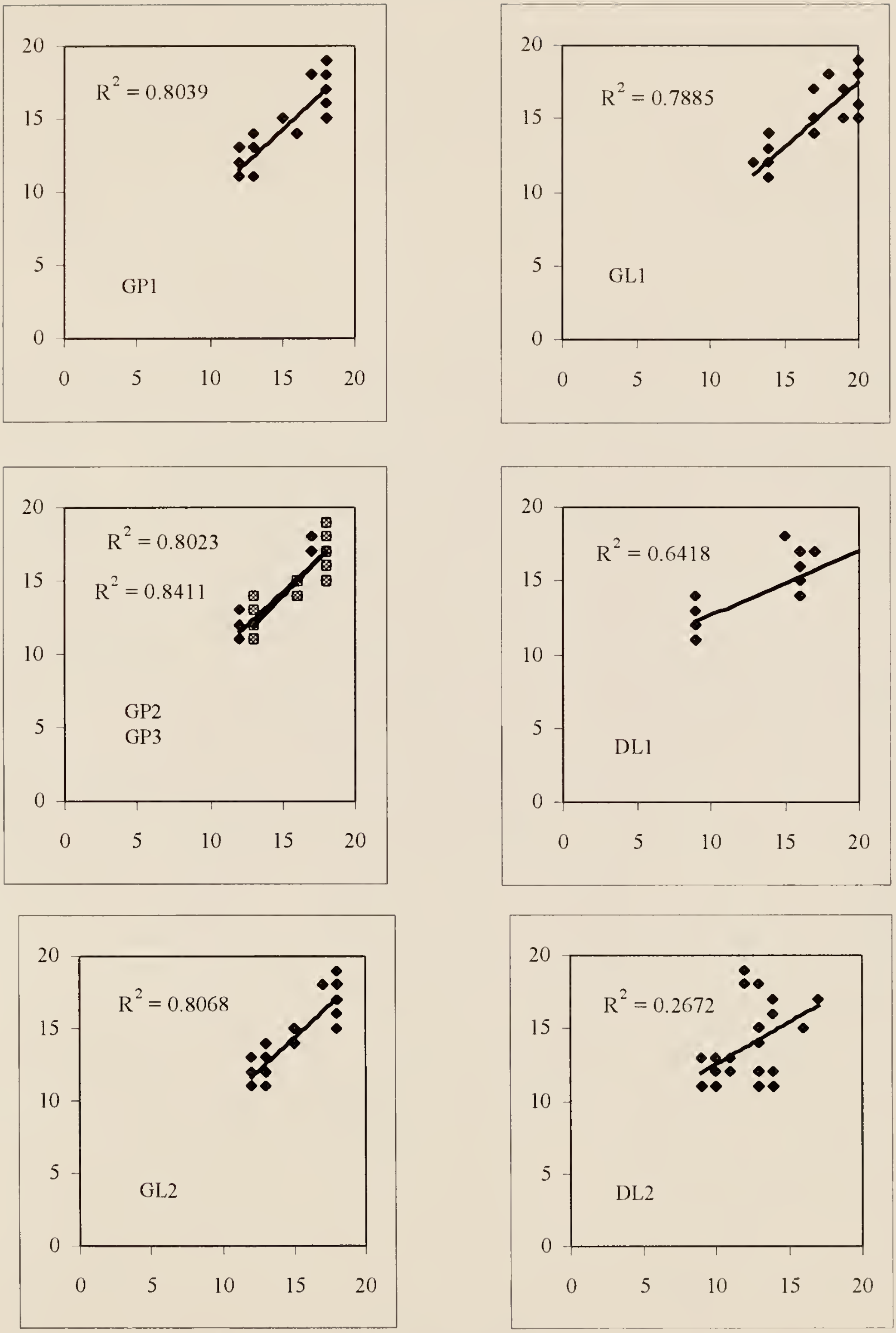
The second arterial street data set containing the new offset design allowed CORSIM and the TRANSYT-7F candidate models to produce (28 phases * 4 programs) 112 data points affected by progression. The existing TRANSYT-7F actuated control model produces 28 data points of its own on these phases, bringing the total to 140, but it is already certain that this oversimplified model is not sensitive to progression effects. There are (21 phases * 4 models) 84 other data points for external links unaffected by progression, or 105 including the existing TRANSYT-7F actuated control model. By comparing the overall internal link results with the external link results, it may be possible to ascertain the effects of progression on actuated phase times. These experiment #2 results will first be presented in tabular format and then in graphical format for better clarity.

Table 4-10 lists the internal link phase times, and table 4-12 lists the external link phase times. Figure 4-6 and figure 4-7 present the same data from tables 4-10 and 4-12 in a different format.

Table 4-10: Offset (Progression) Effects on Internal Link Phase Times

CORSIM	GP1	GL1	GP2	DL1	GL2	DL2	GP3
11	13	14	13	9	13	9	13
12	12	13	12	9	12	11	13
16	18	20	18	16	18	14	18
14	13	14	13	9	13	13	13
11	12	14	12	9	12	13	13
15	18	20	18	23	18	16	18
13	13	14	13	9	13	9	13
19	18	20	18	21	18	12	18
12	12	14	12	9	12	10	13
15	15	17	16	16	15	13	16
12	12	13	12	9	12	10	13
18	17	18	17	15	17	13	18
12	12	14	12	9	12	14	13
17	18	19	18	17	18	17	18
11	13	14	13	9	13	9	13
13	13	14	13	9	12	11	13
17	18	19	18	16	18	14	18
12	12	14	12	9	13	13	13
11	13	14	13	9	12	13	13
15	18	19	18	23	18	16	18
11	13	14	13	9	13	9	13
18	18	20	18	21	18	12	18
11	13	14	13	9	12	10	13
14	16	17	16	16	15	13	16
13	12	14	12	9	12	10	13
18	18	20	18	15	17	13	18
11	12	14	12	9	12	14	13
17	18	17	17	17	18	17	18

Table 4-10 shows a few unexpected individual results for certain individual phases. Overall, correlation with CORSIM still looks better for the models (GXX) that apply the queue service time — green extension time strategy, relative to the models (DXX) that apply the target degree of saturation approach. These results are also



All graphs illustrate actuated phase times in units of seconds
Y-axes contain CORSIM phase times, X-axes contain candidate model phase times

Figure 4-6: Correlation of Phase Times on Internal Links

consistent with those from experiment #1 where the GL1 model computes relatively high green extension times and phase times.

Figure 4-6 shows that the candidate models produce mediocre phase time correlation with CORSIM when internal links are being analyzed, although there is sizable improvement over the existing model. Apparently progression effects do not invalidate the GP3 prototype model, the only candidate model using its own unique queue service time model. Since experiment #1 demonstrated better agreement from the candidate models, correlation losses may be due to differences in the CORSIM and TRANSYT-7F platoon dispersion sub-models, or perhaps a lack of variety with the internal link volumes and link lengths. Table 4-11 lists supplementary candidate model output that indicates the flow profile-based prototype model is reacting appropriately to changes in vehicle arrival patterns.

The table 4-11 results for phase #2 (critical internal link number 1406) show that queue service time decreases by 2 seconds under the improved offset design because more vehicles are arriving on green, resulting in lower average queue lengths. The prototype model (GP1) output shows that the green extension time increases by nearly 2 seconds, producing a total phase time of (13 + 5) 18. Indeed, higher green extension times are expected when progression is good and vehicles are arriving on green. These results are also consistent with the queue service time — green extension time feedback effect observed in experiment #1. Conversely, the NCHRP (GL1) model output for phase #2 shows that the green extension time is not affected by progression. Progression has no impact on green extension times calculated by any of the candidate models besides GP1.

Table 4-11: Numerical Evidence of Progression Effects on Actuated Phase Times

	Critlink	Fq	Gs	Ge	Y+R	Split
	1408	1	7.33	3.21	2	13
GP1	1406	1	12.6	3.35	2	5
Offset Design #1	0	1	0	0	2	54
	1401	1.03	21.8	4.06	2	28
	1408	1	7.32	3.21	2	13
	1406	1	10.6	5.11	2	5
Offset Design #2	0	1	0	0	2	54
	1401	1.03	21.8	4.06	2	28
	1408	1	7.22	4.34	2	14
GL1	1406	1	12.5	4.61	2	5
Offset Design #1	0	1	0	0	2	52
	1401	1.02	21.5	5.11	2	29
	1408	1	7.21	4.34	2	14
	1406	1	10.5	4.61	2	3
Offset Design #2	0	1	0	0	2	54
	1401	1.02	21.5	5.11	2	29

Critlink = TRANSYT-7F critical link number
Fq = queue calibration factor
Gs = queue service time
Ge = green extension time
Split = estimated phase time

External link results

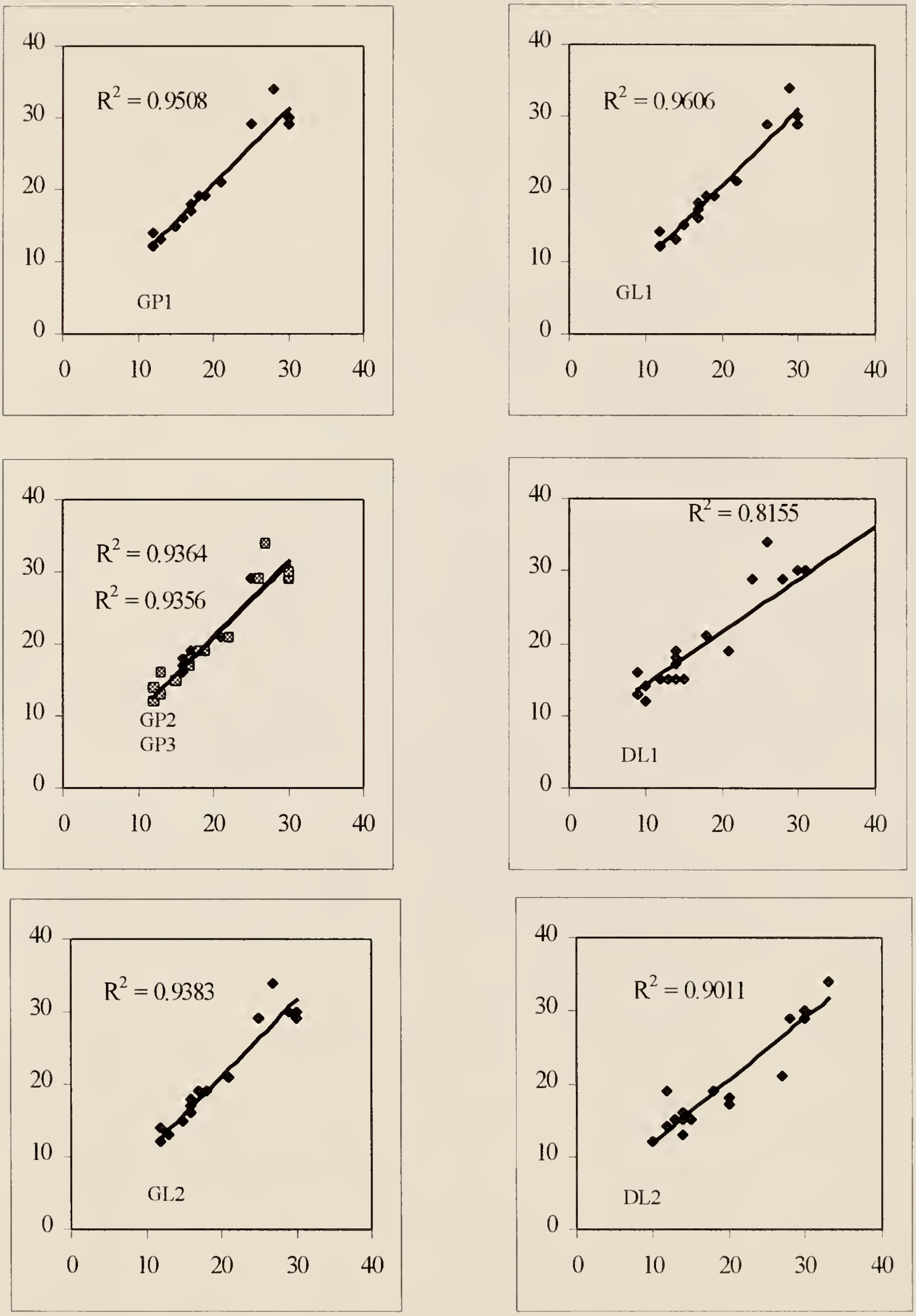
In the absence of queue spillback effects, external link phase times are believed to be completely unaffected by the offsets or progression along the arterial street. The results mostly support this hypothesis, because only a couple of external link phase times changed by one second after switching the offsets, whereas many of the internal link phase times changed. These changes in the external link phase times, resulting from

offset changes, are likely due to inappropriate but negligible biases in the CORSIM and TRANSYT-7F traffic flow models. Table results are shown for one offset design only.

Table 4-12: Offset (Progression) Effects on External Link Phase Times

CORSIM	GP1	GL1	GP2	DL1	GL2	DL2	GP3
19	19	19	18	21	18	12	19
15	15	15	15	14	15	14	15
29	30	30	30	28	30	28	30
15	15	15	15	15	15	15	15
19	18	18	17	14	17	18	18
17	17	17	16	14	16	20	17
18	17	17	16	14	16	20	17
29	25	26	25	24	25	30	26
14	12	12	12	10	12	12	12
15	15	15	15	13	15	15	15
16	16	17	16	9	16	14	13
12	12	12	12	10	12	10	12
15	15	15	15	13	15	13	15
30	30	30	30	30	30	30	30
15	15	15	15	12	15	14	15
30	30	30	30	31	29	30	30
15	15	15	15	12	15	13	15
30	30	30	30	41	30	30	30
34	28	29	27	26	27	33	27
13	13	14	13	9	13	14	13
21	21	22	21	18	21	27	22

Figure 4-7 shows that the GP1 model produces the best phase time correlation with CORSIM when external links are being analyzed; however, this result may be misleading. It was stated earlier that the candidate models sometimes underestimate phase times due to competing link effects. This bias occurs on phases with multiple actuated links that may be critical, because the existing queue calibration factor (f_q) does



All graphs illustrate actuated phase times in units of seconds
Y-axes contain CORSIM phase times, X-axes contain candidate model phase times

Figure 4-7: Correlation of Candidate Model Phase Times on External Links

not take into account volume levels of the competing links. In figure 4-7 there are two data points that are noticeably outliers because competing link effects cause them to be underestimated by the candidate models. In table 4-11, the shaded cells in phase #4 (critical external link 1401) suggest that the GL1 model's penchant for estimating high green extension times is partially counteracting the competing link bias. For this link, CORSIM has estimated an average phase time of 34 seconds, and the inflated green extension time allows the GL1 model (29 seconds) to have better correlation than the GP1 model (28 seconds).

Once the competing link biases are eliminated through use of an improved queue calibration factor, these results should gravitate towards the earlier results of experiment #1. Instead of calculating 1.02 or 1.03, the improved queue calibration factor should be computing a value like 1.20 or 1.30 because on that phase is a competing link with very high volume. Subsequently, the prototype models should produce the best phase time correlation with CORSIM, and the GL1 model should produce slightly overestimated phase times.

Another observation that is available from figures 4-6 and 4-7 is that the candidate models have better correlation with CORSIM on external links (R^2 = 0.95 approximately) than on internal links (R^2 = 0.80 approximately). This may have something to do with differences in the traffic flow sub-models employed by CORSIM and TRANSYT-7F, or it may be caused by the lack of variety associated with the testing data on internal links.

Experiment #2: conclusions

In summary, the experiment #2 testing results indicate that the GP1 model uses its knowledge of the flow profile in order to intelligently compute green extension times. The other candidate models are not far behind in terms of correlation, but do not recognize progression effects as they should.

The external link results also illustrate that the candidate models are capable of producing excellent correlation with CORSIM, likely better than R^2 = 0.97 if the competing link bias can be minimized. The decrease in correlation on internal links is possibly caused by differences in the overall CORSIM and TRANSYT-7F models, or lack of variety with the internal link testing data. Another factor that will be shown to affect correlation between CORSIM and TRANSYT-7F is the modeling of permitted left-turns.

Experiment #2 Supplemental, Signal Timing Optimization Comparison

In chapter 1 it was stated that the optimization capabilities of TRANSYT-7F should be improved as a result of the improved simulation capabilities. The methodologies within this dissertation directly improve the accuracy of TRANSYT-7F simulation through improvements to the accuracy of actuated phase time calculation. The methodologies also indirectly improve TRANSYT-7F optimization, because the optimization is simply a series of simulation runs, in which the simulation run that produced the best results is reported as optimal.

Experiment #2 supplemental was conducted to demonstrate an example of the potential improved optimization. Two copies of the arterial street benchmark data set (used to generate the experiment #2 results) were used in this experiment. Both data sets

were modified so as to request optimization instead of simulation. In one of the data sets, optimization was performed using the existing TRANSYT-7F release 8 program. In the other data set, optimization was performed using the GP1 candidate model. The optimization objective function used in both cases was the standard TRANSYT-7F disutility index (DI), which is a combination of vehicle delay and stops.

The two sets of optimal timing plans were then coded inside two separate CORSIM input files, in order to evaluate those timing plans. The CORSIM input files were also copies of the arterial street benchmark data set used to generate the results from experiment #2. Because actuated phase times cannot directly be optimized, force-offs were not modified from those used in experiment #2. Only the two sets of optimal offsets resulting from the "old" and "new" TRANSYT-7F actuated control methodology were exported (and defined as yield points) into the CORSIM data sets. Because the reference intervals defined in TRANSYT-7F were the yellow intervals immediately following the non-actuated green, no translation was necessary before coding these offsets into CORSIM as yield points. The two designs of offsets or yield points, along with the corresponding simulation results from CORSIM, are listed in table 4-13.

Table 4-13 indicates an approximate 5% reduction in average delay per vehicle, using the optimal offsets based on the new actuated control methodology. The data from experiment #2 is ideal for testing optimization since there are few complications, such as queue spillback or permitted left-turns, which could introduce a bias into the results. While it is difficult to ascertain the amount of delay reduction (under the new actuated control methodology) to be expected on most traffic networks, it is clear that optimization results are predicated on getting the correct results from deterministic simulation. The

methodologies from the study are specifically designed to improve accuracy of the results from deterministic simulation.

Table 4-13: Optimization Comparison under Old and New Methodologies

Offset design generated by:			CORSIM output based on timing from:		
	Old	New		Old	New
Node	T7F	T7F		T7F	T7F
6	0	50	Control Delay	25.6	24.2
7	15	99	Total Delay	29.7	28.4
8	29	16	Stop Delay	22.0	20.5
9	41	20	Vehicle Trips	8116	8242
10	53	10			
11	27	99			
12	27	27			
13	57	74			
14	18	76			
15	64	64			

Experiment #3, Permitted Left-Turn Effects

In experiment #2, two data sets were created with different offset designs in order to test progression effects on actuated phase times. Both of these data sets had protected-only actuated left-turn phases, in which only green left-turn arrows would be displayed by the signal in the field. Typically permitted-only left-turn phases, in which only a green ball is displayed by the signal in the field, are not actuated because the length of the phase is designed to handle through movement traffic. However, protected-permitted left-turn phases, in which a green arrow followed by a green ball is displayed in the field, are frequently actuated.

It is more complicated to estimate actuated phase times on protected-permitted phases because the results are totally dependent on how much traffic is served during the

permitted portion of the phase. If the opposing movement has heavy traffic and few left-turns are served during the permitted portion of the phase, then the actuated phase time may be nearly equal to what it would be under protected-only phasing. If the opposing movement has light traffic and many left-turns are served during the permitted portion, then the actuated phase time may be nearly equal to the minimum phase time, because the queues are so small during the protected portion. The queue accumulation polygon (QAP) illustrated in figure 4-8 shows that the queue length accumulated during the permitted portion directly affects the subsequent queue service time.

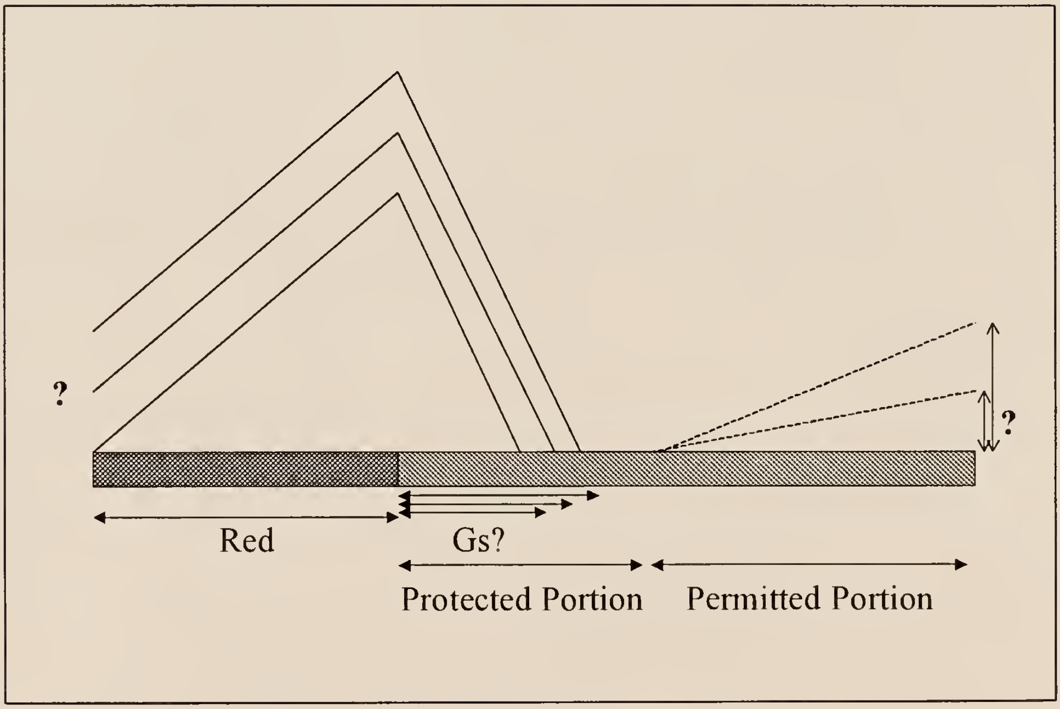


Figure 4-8: Multiple QAP Possibilities for Protected-Permitted Phasing

The existing model within TRANSYT-7F for calculating actuated phase times is inadequate because it does not attempt to estimate phase times for protected-permitted

actuated phases. Instead, such phases are currently set to their minimum phase times automatically, as if all traffic was served during the permitted portion of the phase. It would be difficult to rectify this model in its existing form because the amount of permitted portion traffic served is determined during simulation, after the phase time has already been computed by the existing model. The candidate models elegantly avoid this pitfall because TRANSYT-7F automatically tells them when the queue has been served, which is automatically affected by how many vehicles were served during the permitted portion of the phase.

Testing conditions: experiment #3 for evaluating permitted left-turn effects

The purpose of experiment #3 is to ascertain permitted left-turn effects on actuated phase times. In this experiment, protected-permitted phasing was defined on numerous left-turn links within the original arterial street data sets for CORSIM and TRANSYT-7F. However, one additional CORSIM data set was created specifically to achieve better consistency with TRANSYT-7F. The reason for this is that CORSIM contains a couple of default parameter values that would definitely bias the results and obscure the permitted left-turn effects on actuated phase times. By default, on permitted left-turn links, CORSIM implements 0.38 "jumpers" per cycle, plus a statistical distribution for "sneakers". However, in the benchmark data set, TRANSYT-7F implements 0 jumpers and 2 sneakers per cycle. Thus the second data set, whose results are labeled "Modified CORSIM" in experiment #3, contains modified input data that attempts to achieve 0 jumpers and 2 sneakers per cycle, similar to TRANSYT-7F.

The experiment #3 results will first be presented in tabular format and then in graphical format for better clarity. Certain phases are totally unaffected by permitted left-

turn effects, such as a protected actuated phase that occurs "earlier" in the cycle than any permitted phase. This terminology assumes that a cycle "begins" at the end of the non-actuated phase. Chapter 3 observations on early return to green effects demonstrate that actuated phases are sometimes affected by the performance of other actuated phases that occur earlier in the cycle. Thus, protected phases that occur after any permitted phase are listed along with protected-permitted phases in the cells of table 4-14, resulting in (30 * 6) 180 data points for testing. Figure 4-9 and figure 4-10 present the same data from table 4-14 in a different format.

Table 4-14 shows a few unexpected individual results for certain individual phases; however, correlation with CORSIM still looks reasonable for the candidate models. Correlation is quite poor for the existing model within TRANSYT-7F (DL1) because it has no methodology for adjusting the target degree of saturation in response to permitted left-turns. The DL2 model also has no methodology for this, and so it was omitted from experiment #3 altogether. The GL1 model results are consistent with those from experiments #1 & 2, which indicate that this model has a tendency to compute relatively high green extension times and phase times.

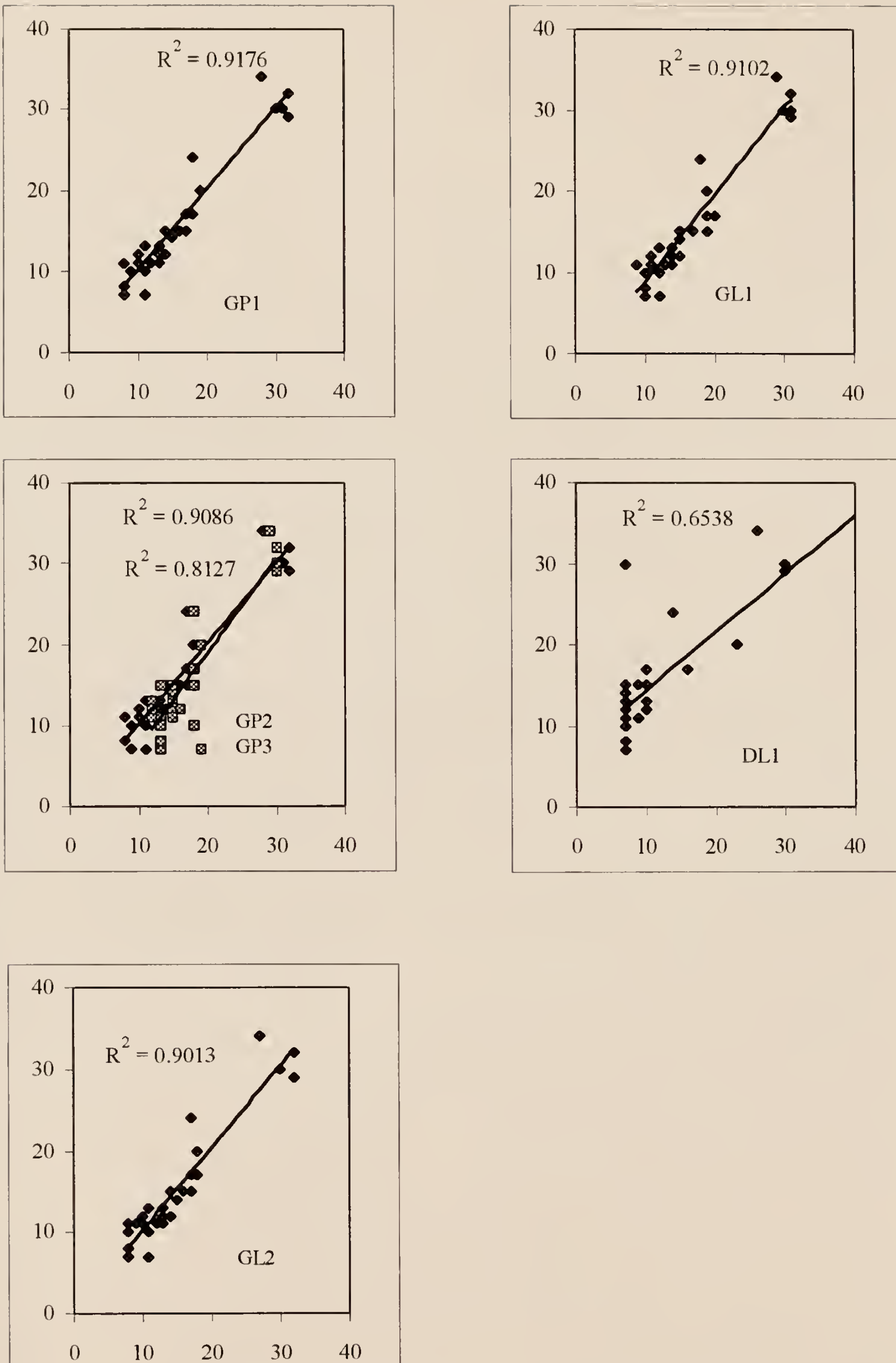
Experiment #3 results and conclusions

Figure 4-9 shows that the GP1 model produces the best correlation in the permitted link network, although the other candidate models are close behind. Correlation of the existing model within TRANSYT-7F (DL1) is relatively poor for the permitted link network as expected. In addition, GP3 model performance deteriorates, because the uniform queue service time computation described in chapter 2 cannot react to the complexity of permitted left-turns. Figure 4-10 shows that candidate model

correlation improves after calibrating CORSIM so it would simulate permitted left-turns having 0 jumpers and approximately 2 sneakers per cycle, similar to TRANSYT-7F.

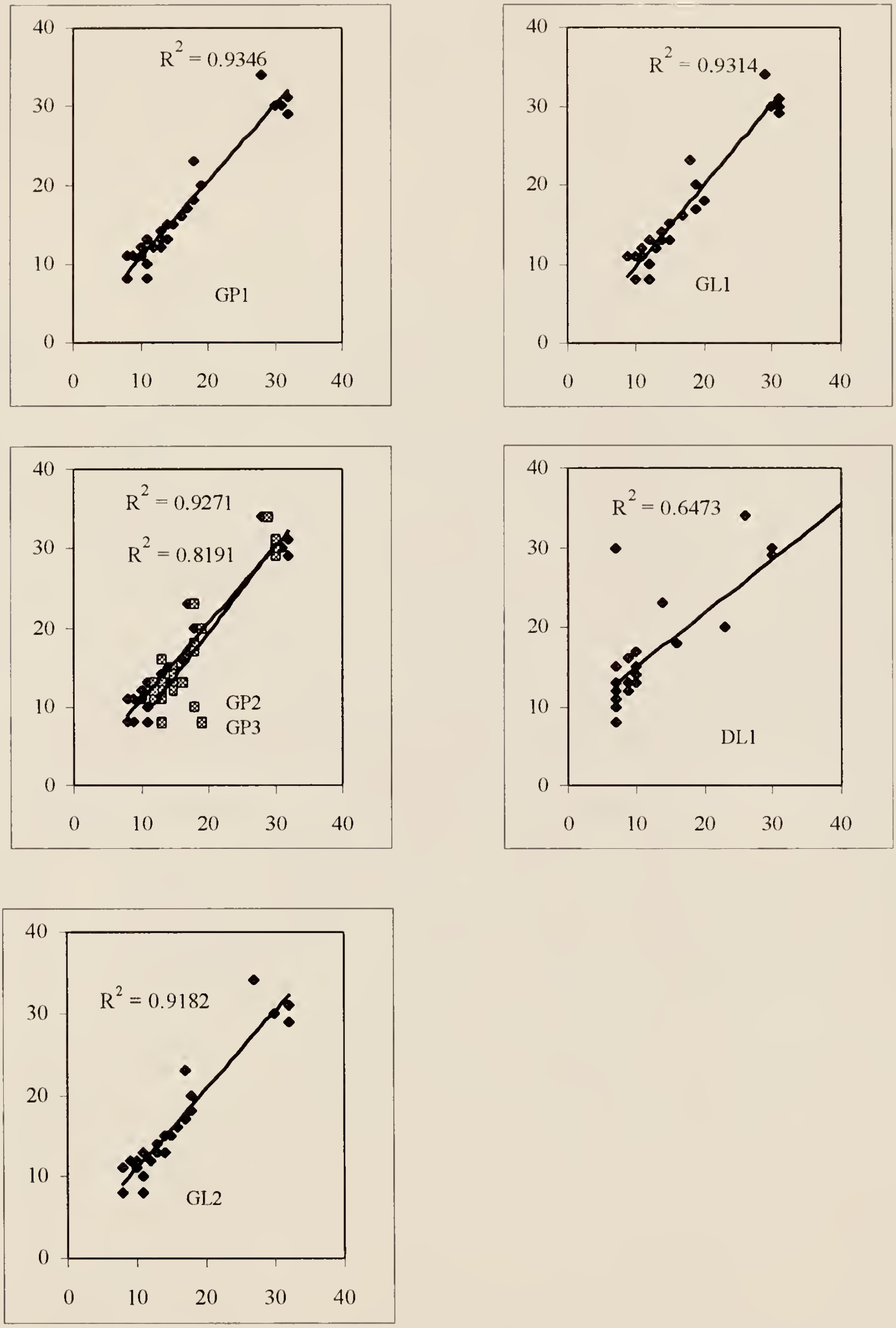
Table 4-14: Permitted Left-turn Effects on Phase Times (Sorted by Model)

CORSIM	CORSIM Modified	GP1	GL1	GP2	DL1	GP3	GL2
11	13	13	14	13	9	13	13
20	20	19	19	18	23	19	18
15	15	14	15	14	7	15	14
30	30	31	30	31	30	30	30
11	12	13	13	12	9	13	12
17	18	18	20	18	16	18	18
14	15	15	15	15	7	15	15
24	23	18	18	17	14	18	17
11	11	8	9	8	7	13	8
7	8	8	10	9	7	19	8
13	13	11	12	11	7	12	11
15	15	14	15	14	10	15	14
15	16	16	17	16	9	13	16
11	12	10	11	10	7	13	9
15	17	17	19	17	10	18	17
11	11	10	11	10	7	12	10
12	13	13	14	13	10	15	13
29	29	32	31	32	30	30	32
7	8	11	12	11	7	13	11
12	13	14	15	14	10	16	14
11	12	12	13	12	7	15	12
30	30	30	31	30	7	30	30
10	10	11	12	11	7	18	11
12	12	10	11	10	7	12	10
13	14	13	14	13	10	15	13
32	31	32	31	32	41	30	32
10	11	9	10	9	7	13	8
17	17	17	19	17	10	18	17
34	34	28	29	28	26	29	27
8	8	8	10	8	7	13	8



All graphs illustrate actuated phase times in units of seconds
Y-axes contain CORSIM phase times, X-axes contain candidate model phase times

Figure 4-9: Phase Time Correlation Assuming CORSIM's Perm. Left-turn Defaults



All graphs illustrate actuated phase times in units of seconds
Y-axes contain CORSIM phase times, X-axes contain candidate model phase times

Figure 4-10: Phase Time Correlation Assuming CORSIM's Left-turn Modifications

The purpose of this experiment was to demonstrate that the candidate models are able to explicitly recognize permitted left-turns. The test data indicates that reasonable correlation with CORSIM was maintained, even though the microscopic simulation of CORSIM and the flow profile simulation of TRANSYT-7F are likely to predict permitted left-turn capacities that are not identical. The candidate models based on target degree of saturation do not account for permitted left-turns in their existing form. The other candidate models recognize permitted left-turns elegantly because they are predicated on queue service time, and TRANSYT-7F automatically calculates the queue length during each step of simulation. In addition, because of this queue service time structure, the candidate models are automatically able to adjust actuated phase times in response to queue spillback.

Experiment #4, Spillback Effects

When the queue from a downstream link becomes sufficiently long due to congestion, upstream links will experience the blockage that is caused by queue spillback effects. This means that queues at the upstream links will be either temporarily or indefinitely prevented from discharging, no matter how long they are. The structure of the candidate models for actuated control within TRANSYT-7F indicates that actuated phase lengths are fundamentally affected by queue service time. Queue service time itself, however, is fundamentally affected by spillback effects. Thus by transitive properties, actuated phase lengths are fundamentally affected by spillback effects. Figure 4-11 uses CORSIM animation to illustrate queue spillback along the east-west (left to right) arterial street that is currently preventing the queue from discharging at an actuated minor street left-turn phase.

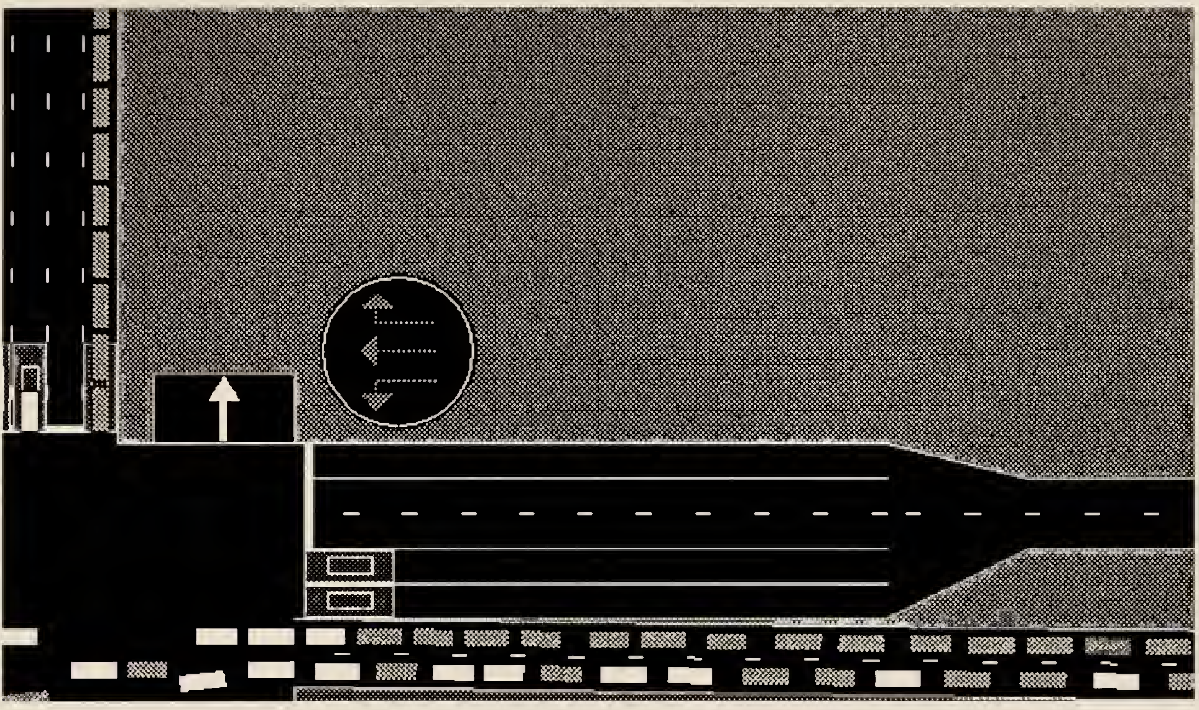


Figure 4-11: Animation Showing Spillback Effects on Actuated Phase Times

As with permitted left-turn effects, the candidate models are able to recognize spillback effects elegantly because TRANSYT-7F automatically calculates the queue length during each step of simulation. One way to test this phenomenon is by generating a test data set having sufficient congestion along the arterial street. Theoretically, optimizing the offsets should alleviate spillback along the arterial street, which should subsequently alter the actuated phase times on links that were experiencing blockage. This raises the question about where the candidate models fit within the context of signal timing optimization. Thus, an experiment for ascertaining spillback effects will be combined with an optimization experiment.

Optimization Applications Testing

Traditional signal timing strategies involved optimizing the splits (green times). However, with the advent of traffic-actuated control, a maximum green time (or force-off), a minimum green time, and a gap setting are now sometimes specified instead of the green split. The result is that the green split cannot always directly be optimized.

The candidate models presented in this chapter are actually designed to estimate the actuated green split that can no longer be optimized. Why is it important to estimate the average green splits when instead the focus of optimization programs like TRANSYT-7F should be on optimization? The reason is that optimization programs (and capacity analysis programs) need to know the correct, average amount of green time available on the arterial street. Subsequently, optimization programs are able to intelligently generate a design that can achieve progression. When the correct amount of green time available to the arterial street is known, this allows for subsequent optimization of the phasing pattern, cycle length, and offsets, although TRANSYT-7F itself does not currently optimize phasing pattern.

New Optimization Paradigm

So proper optimization depends on knowing the correct actuated phase times, but in fact the phase times are likewise affected by the same signal settings (phasing pattern, cycle length, offset) that can be optimized. The result is a chicken-and-egg situation between phase time calculation and signal setting optimization. The existing model for actuated control within TRANSYT-7F does not account for this circular dependency, because it produces the same phase time estimates regardless of the optimal signal

settings. However, because the candidate models are appropriately responsive to network-wide queue lengths and flow profiles, it should now be possible to thoroughly analyze this circular dependency for better results.

Theoretically, if a continuous series of phase time calibration runs and signal setting optimization runs converges to produce one unique network-wide signal timing plan, a new level of accuracy and optimality should be possible. This iterate-until-convergence concept is similar to the candidate models' strategy for calculating phase times alone. Because the length of any actuated phase depends on the length of all other actuated phases, individual phase times should be continuously recalculated in response to one another until all are in agreement. Convergence is clearly achieved when additional recalculations do not change the phase times whatsoever. Subsequently the same concept can be applied to the bigger picture. Because the actuated phase lengths and optimal signal settings depend on each other, they should be continuously recalculated and re-optimized until all are in agreement. Convergence is clearly achieved when additional recalculations do not change the phase times whatsoever, and additional re-optimizations do not improve network performance whatsoever. Figure 4-12 illustrates this phenomenon.

Experiment #4 will be used as an optimization case study in which spillback effects will also be observed. In this experiment, phasing pattern, cycle length, and maximum greens (force-offs) are held constant, and the focus is on actuated phase time calibration and offset optimization. The optimization objective function applied was the traditional TRANSYT-7F disutility index (DI), which represents a combination of vehicle delay and stops. Traffic volumes were doubled in the benchmark data set from

experiment #2 to create a new data set with oversaturated conditions for experiment #4, whose results are listed in table 4-15.

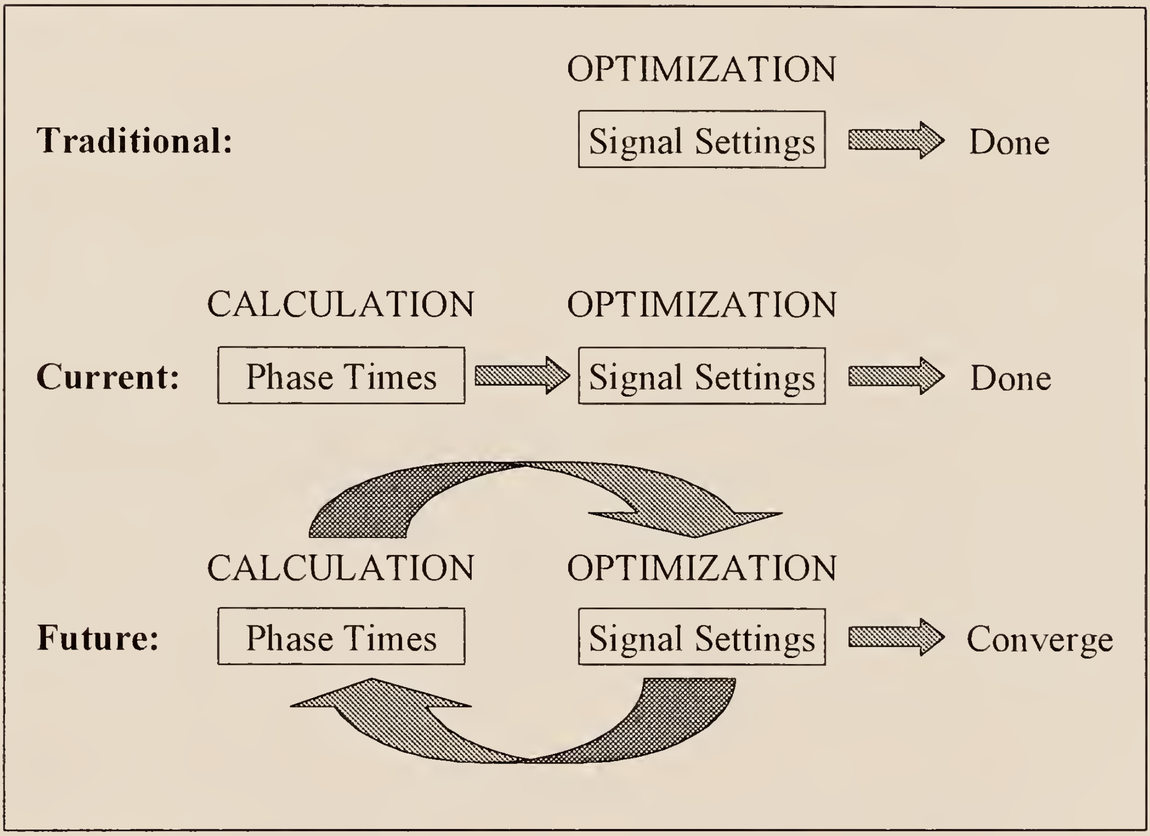


Figure 4-12: Evolution of Signal Timing Optimization Strategies

Although this network used for testing is considered oversaturated, some of the individual actuated phases in fact remained undersaturated, even after doubling traffic volumes from the original benchmark data set. Because of this, the updated actuated control model was successfully able to reassign unused green time to the oversaturated phases, which subsequently reduced spillback and vehicle delay.

Table 4-15: Experiment #4 Optimization Results

	Phase Times						Offset	Delay	Flow
Before	22	3	30	15	30		37	942	
Optimization	22	3	30	15	30		56		
(GP1)	60	20	20				16		
	15	55	30				35		
	22	3	30	12	3	30	55		3720
	22	3	30	12	3	30	14		
	22	3	30	15	30		33		
	22	3	30	15	30		53		
	22	3	35	40			12		
	15	55	30				32		
First	16	9	30	15	30		67	579	
Optimization	17	3	38	15	27		2		
	60	20	20				1		
	15	55	30				52		
	17	3	38	12	3	27	51		3933
	17	3	35	12	3	30	25		
	17	3	35	15	30		48		
	15	4	36	15	30		87		
	18	3	39	40			29		
	15	55	30				49		
Ninth	16	9	30	15	30		96	582	
Optimization	17	3	38	15	27		94		
	60	20	20				92		
	15	55	30				37		
	17	3	43	12	3	22	36		4013
	17	3	35	12	3	30	12		
	17	3	35	15	30		40		
	17	3	35	15	30		83		
	18	3	39	40			39		
	15	55	30				58		
Existing	9	12	37	14	28		90	460	
T7F	11	9	43	19	18		86		
(DL1)	66	17	17				22		
	11	56	33				14		
	11	19	42	13	3	12	65		4660
	9	13	33	10	3	32	50		
	10	8	31	14	37		40		
	10	6	26	13	45		97		
	13	12	35	40			39		
	15	52	33				58		

Table 4-15 contains four quadrants of results. The first quadrant on the top lists the phase times, offsets, and vehicle delay for the test network prior to actuated phase

time calibration and offset optimization. For this experiment, actuated phase times were calculated using the GP1 model. The second quadrant lists the same results following the first iteration of phase time calibration and offset optimization. These results show that the GP1 model has readjusted many of the phase times, TRANSYT-7F has optimized many of the offsets, and network-wide delay has decreased as expected.

Progression Effects on Vehicle Delay, Phase Times, and Spillback

Attempting to implement the iterative procedure described in figure 4-12, phase times and offsets were continuously recalculated until network-wide delay converged at 582 seconds per vehicle. The delay observed at each iteration was the following: 942, 579, 595, 578, 581, 582, 577, 582, 582, 582. After the ninth iteration, delay is no longer changing and the process has converged. The third quadrant of results reveals that only 3 out of 35 actuated phase times have changed following the initial calibration run, although all of the offsets have changed. Most actuated phase times on external links are not changing because the number of vehicles arriving is more than can be served, and these phases are being forced off by the maximum green. Most actuated phase times on internal links are not changing. The queue service time — green extension time feedback effect may be counteracting progression effects caused by the new offsets.

The shaded cells indicate three actuated phase times that did change during the optimization process. At intersection #8, two of the internal link left-turn phases have shifted due to progression effects. However, the external link right-turn phase time has decreased noticeably. This happens because the optimal offsets have alleviated spillback along the arterial street, thus allowing the minor street right-turn queue to be served more efficiently.

Although the network-wide delay did not change significantly after the first iteration of optimization, the output called "Flow" indicates that subsequent optimizations are increasing capacity. This output represents the number of vehicles that are discharged. Initially, intersection #10 is getting blocked by upstream spillback, and only 3720 vehicles are processed by the intersection. After the first optimization, arterial street spillback is alleviated to some extent, which allows 3933 vehicles to the processed at the same intersection. Subsequent optimizations that further eliminate spillback also increase the efficiency of the minor street right-turn. Network-wide delay is not reduced, since better service to the minor street right-turn simply supplies additional congestion to the arterial street. However, 4013 vehicles are now being processed by intersection #10, which indicates a more efficient operation.

Existing Models: Overly Optimistic Results

By contrast, the existing model within TRANSYT-7F seriously underestimates the minor street right-turn phase time, because spillback effects are not recognized. Consequently, an unrealistically high amount of green time is made available to the arterial street. Ironically, this extra arterial green time significantly alleviates spillback, allowing 4660 vehicles to be served at intersection #10. The real danger here is that this optimal design, under the existing model, will never materialize in the field. Because actuated phase times will be longer than those calculated by the existing model, program outputs like delay (460) and flow (4660) are inappropriately optimistic.

On undersaturated phases, the existing (DL1) model calculates phase times that are usually lower than those from the GP1 model, resulting in more green time available for the congested phases. On oversaturated phases, the existing model ignores the

maximum green, again resulting in more green time available for the congested phases. The end result of applying the existing model is a solution that is overly optimistic, and less likely to materialize in the field, relative to designs generated with assistance from the candidate models. Overly optimistic results such as these are possible with models that tend to underestimate phase times, and with models that do not recognize spillback effects. However, these results do reveal that better system performance would be attainable by implementing lower maximum green times that would forcibly lower the actuated phase times.

Experiment #4 Summary

First, external link phase times should be totally unaffected by offsets or progression effects, so a decreasing external phase time is an indicator of spillback effects being alleviated by optimal offsets. Secondly, since the first iteration of phase time calibration and offset optimization was not sufficient in reducing spillback or achieving delay convergence, this indicates codependency of the offsets and actuated phase times, which is not appropriately handled by the traditional single-step optimization procedures from figure 4-12. Finally, since network delay is unchanged as the number of processed vehicles increases, while phase times and offsets are shifting, this indicates a typical effect of signal timing optimization associated with oversaturated conditions and spillback effects. The huge shifts in the offsets indicate that similar performance is possible with a variety of offset designs, whereas the minor shifts in the phase times indicates that they are becoming more accurate and realistic.

Chapter Four Summary

In this chapter experiments were performed using hypothetical test data representing a single intersection first, and then an arterial street. The initial simple experiments using a single intersection resulted indicated that candidate models applying the queue service time — green extension time strategy were producing better correlation with CORSIM, relative to their counterparts applying the target degree of saturation strategy. The single intersection experiments also revealed the concept of queue service time — green extension time feedback effect, and showed some preliminary differences between the candidate models.

Testing was performed on arterial street data in order to ascertain progression effects, permitted left-turn effects, and spillback effects. During these tests, the queue calibration factor was deemed to be useful in minimizing competing link effects, but also in need of updating in order to improve correlation with stochastic programs like CORSIM. Regarding progression effects, all models were shown to produce mediocre phase time correlation with CORSIM on internal links. The GP1 model was shown to produce logical green extension times in based on platooned vehicle arrivals. Regarding permitted left-turn effects, the candidate models were shown to recognize this explicitly and produce reasonable correlation with CORSIM, given that it tends to generate different permitted left-turn capacities than TRANSYT-7F. Regarding spillback effects, the candidate models were said to recognize this explicitly because, as with permitted left-turns, queue service time can simply be taken from the results of TRANSYT-7F

simulation. Also, evidence of spillback effects was cited during an experiment on signal timing optimization.

The combined body of testing data, compiled and presented in chapter 4, was also consolidated and used to demonstrate overall improvement due to the methodology within this dissertation. These results were presented on two graphs in chapter 1. These graphs compare actuated phase times computed by the existing model within TRANSYT-7F (DL1), and the candidate model (GP1) that showed the best performance in the chapter 4 testing results. The test results were obtained from experiments #1a, 2 and 3. They did not include results from experiment #4 because CORSIM was not use to analyze the oversaturated testing conditions. They also did not include results from experiment #1b because the DL1 model is not capable of analyzing fully-actuated control. Capability to analyze fully-actuated control is yet another advantage of the candidate models that apply the queue service times — green extension time strategy.

After testing the candidate models as mentioned above, the new models were hypothesized to change the way signal timing optimization can be performed now, since the effect of optimal signal settings on actuated phase times can now be recognized. A case study experiment was performed to examine how the candidate models would fit in to the context of signal timing optimization. The results of this experiment suggested that the codependency of actuated phase times and optimal signal settings can indeed be recognized, and that continuous iterations of phase time calibration and signal setting optimization can result in better timing plans, in addition to improved phase time accuracy.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

In this chapter, conclusions are drawn primarily on the results of chapter 4 testing, and to a lesser extent, based on theoretical discussion from chapters 2 and 3. Subsequently, recommendations are given regarding modeling and future research. The initial conclusions involve evaluating the candidate models for improvement of TRANSYT-7F performance.

Conclusions

To review, the candidate models are designed to estimate average phase times under traffic-actuated control. The primary difference between the candidate models lies in the calculation of green extension time, after the queue has been served. Except for the GP3 model, each candidate model obtains queue service times directly from the results of TRANSYT-7F step-wise simulation. At isolated intersections, queue service times from simulation are typically identical to those estimated by the NCHRP queue service time formula. However, inside signalized networks, queue service times

are affected by adjacent intersections, and thus should be obtained from simulation instead of the formula. The GP3 model uses uniform queue service times, which are calculated according to the steps outlined in chapter 3.

Candidate Model Performance

First, an ideal candidate model would have a good theoretical background or base. Models having solid theoretical foundations are less likely to break down under unexpected input conditions, tend to be easier to understand, and facilitate future enhancements or development. Second, an ideal candidate model would have a minimal number of required calculations, in order to minimize program running times and programming errors. Finally, an ideal candidate model would display excellent correlation with the phase times simulated by CORSIM, the detailed traffic simulation program. Table 5-1 is an attempt to convey perceived performance of the candidate models, based on these evaluation criteria.

In this table, scores of 1 (bad), 2 (fair), and 3 (good) were assigned to each of the candidate models for multiple evaluation categories. These scores are subjective interpretations of performance under chapter 4 testing, and theoretical background from chapters 2 and 3. For example, the GP3 model receives a score of 1 (bad) for experiment #4 because it cannot handle spillback effects. Models with the highest overall score on the bottom of the table are concluded to have the best overall performance. Models that apply the target degree of saturation strategy were omitted from this table due to their shortcomings, and due to their lack of correlation with the other models.

Table 5-1: Candidate Model Evaluation Scorecard

Evaluation Criteria	Candidate Model				
	GP1	GL1	GP3	GP2	GL2
Theoretical Foundation					
Isolated Intersections	2	3	2	2	2
Coordinated Intersections	3	2	1	1	1
Computational Speed & Ease of Implementation	1	2	3	2	2
Correlation with CORSIM					
Experiment #1a	3	2	3	1	1
Experiment #1b	3	1	2	1	1
Experiment #2-internal	2	2	3	2	2
Experiment #2-external	3	3	2	2	1
Experiment #3	3	3	1	3	3
Experiment #4	3	3	1	3	3
Score:	23	21	18	17	16

Figure 1-2 from chapter 1 illustrates the improved candidate model correlation of actuated phase times with CORSIM, relative to the mediocre correlation that was evident from the existing model in figure 1-1 from chapter 1. The situation is even more optimistic than these figures indicate. First of all, unlike the existing model, the candidate models are appropriately responsive to spillback effects, and figures 1-1 and 1-2 from chapter 1 contain no spillback data. Secondly, unlike the existing model, the candidate models are capable of computing actuated phase times for fully-actuated signals. Third, development of an improved mechanism for minimizing competing link effects should further improve correlation. The three data points (outliers) from figure 1-2 that possess the worst correlation would be corrected by an improved queue calibration factor, or calculation of green extension time based on multiple link flow rates. Finally,

unlike the existing model, the candidate models are sensitive to optimization effects, thus allowing the new optimization paradigm described in chapters 3 & 4.

The Achilles heel of the GP1 model may be the fact that it does not explicitly recognize some of the parameters found within the GL1 model, such as number of lanes, detector length, approach speed, and vehicle length. For example, although the GP1 model does divide volume by the number of lanes to obtain a "per lane" flow rate, gap-out behavior for 300 vph across two lanes ought to be slightly different than 150 vph within one lane. On the other hand, the GL1 model has special coefficients for recognizing multilane scenarios, which is more comprehensive treatment than simply obtaining the flow rate per lane. However, the phase time biases caused by these parameters are perhaps being minimized by the TRANSYT-7F traffic flow model, and the queue service time — green extension time feedback effect described earlier. The flow rates and queue lengths simulated by TRANSYT-7F implicitly account for parameters like number of lanes, approach speed, and vehicle length.

Even the most effective candidate models did not have perfect correlation with CORSIM. Based on the testing results, it appears that the biggest remaining obstacles to achieving better correlation are:

1. Differences in the permitted left-turn sub-models employed by CORSIM and TRANSYT-7F
2. Competing link effects

Reconciliation of differences in the permitted left-turn sub-models is beyond the scope of this study. However, a recommendation on how to minimize competing link effects with future research will be given later in this chapter.

Testing Limitations

Testing limitations were listed in chapter 4. Some of these testing limitations are inconsequential, whereas some of these limitations deserve to be scrutinized by future research. The limitations that require more scrutiny are discussed later on in this section regarding recommendations. For now, the following testing limitations should prove inconsequential to the chapter 4 testing results, and to the chapter 5 conclusions and recommendations:

Shared lanes served by permitted, actuated phases

Although this condition was not tested, the candidate models should be able to handle this because of their methodology. In obtaining queue service times from simulation, it doesn't matter whether a link is permitted or protected. Green extension times should not be computed less accurately either. Having said this, conventional traffic engineering wisdom dictates that permitted-only links should rarely be actuated, and shared lanes should rarely be served by protected-permitted phasing.

A grid, or a network of signals

Chapter 4 testing was only performed on single intersections and arterial streets. However, testing results are still applicable to grids or networks of signals. Specifically, external link results from experiments #1-4 are just as valid for network external links, and internal link results from experiments #2-4 are just as valid for network internal links. Candidate model queue service times and green extension times are affected by actuated control parameters, queues, and flow profiles on networks just as they are on single intersections or arterial streets.

Intersections with 5 or more approaches

Just as networks do not adversely affect the candidate models, the number of intersection approaches does not adversely affect the candidate models. Candidate model results are based on actuated control parameters, queues, and flow profiles.

Volume-density controllers, or controllers with gap reduction functionality

Gap reduction features are rarely enabled within coordinated signal systems [Orcutt, 1993], and coordinated signal systems are the primary scope of TRANSYT-7F. In addition, volume-density control may even be uncommon at isolated intersections in many cities. Because of this, the lack of volume-density or gap reduction testing should be inconsequential. However, the candidate models designed for basic actuated control are bound to be less accurate when applied to volume-density control. If volume-density control becomes popular or is popular in a certain jurisdiction, the candidate models may be less effective.

Multi-cycle or multi-period simulation

Multi-cycle simulation is generally applicable for oversaturated networks, because on undersaturated networks, queues will remain unchanged between cycles. On oversaturated networks, the number of cycles that are simulated directly affects the severity of queuing and spillback. Candidate models that obtain queue service times directly from simulation should automatically recognize the level of queuing and spillback, and should thus automatically recognize multi-cycle simulation effects. Although the program that implements the candidate models does not currently handle multi-cycle simulation, it could be updated to do so.

Multi-period simulation is generally used to model 15-minute variations in traffic flow demand. Variation in traffic flow demand can affect network-wide queues and flow profiles; therefore, candidate models that recognize the queues and flow profiles should automatically recognize multi-period simulation effects. Although the program that implements the candidate models does not currently handle multi-period simulation, it could be updated to do so.

Actuated Control Behavior

Chapter 3 described a theory about queue service time — green extension time feedback effect, which was subsequently supported by the results from chapter 4 (experiment #1a,b). The theory states that there is a positive feedback effect under isolated intersection conditions, and a negative feedback effect when a background cycle length is enforced. There are a couple of conclusions to be drawn from the queue service time — green extension time feedback theory:

1. Actuated phase time calculator models that are effective under coordinated conditions may not be as effective under isolated conditions, and vice versa.
2. Inaccurate green extension time estimates are more likely to cause inaccurate actuated phase time estimates under isolated conditions. The positive feedback effect causes models to aggravate the mistake by predicting highly inaccurate queue service times. This is not to say that accurate green extension time estimation is not important under coordinated conditions. Under coordinated conditions, the negative feedback effect will *partially* compensate for errors in green extension time estimation, but will not *fully* compensate for such errors.

Summary of New Modeling Capabilities

The candidate models discussed in this paper should allow for better evaluation and optimization of traffic-actuated signal systems using TRANSYT-7F. In the context of evaluation, program output related to vehicle delay, stops, queuing, etc. become automatically more accurate when actuated phase time estimation is improved. In the context of signal timing optimization, better designs are possible when the correct amount of green time available to the major street is known. The designs are not overly optimistic, nor overly pessimistic, and are more likely to materialize when implemented in the field. Moreover, the candidate models' recognition of network-wide behavior results in a new optimization paradigm. Offsets and actuated phase times can now be continuously optimized and calibrated until the entire network achieves convergence in the form of a balanced signal timing plan.

Inherent aspects of the TRANSYT-7F traffic flow model, combined with new candidate model technology, now have the appropriate impact on program results. The new models recognize actuated control parameters that were previously ignored by the earlier model predicated on target degree of saturation. Extraction of queue service times directly from the results of step-wise simulation, plus prediction of green extension times based on signal settings and the simulated flow profile, allows actuated phase times to recognize numerous complexities including:

- Early Return to Green Effects
- Overlap Phasing Effects
- Stochastic Effects

- Progression Effects
- Permitted Left-turn Effects
- Spillback Effects
- Optimization Effects

Recommendations

Regarding TRANSYT-7F, it is recommended that the GP1 model be implemented as the default actuated phase time calculator in upcoming versions of the program. However, the GL1 model, and possibly other candidates should also be available for selection within the program.

Regarding other deterministic models, including optimization models and the HCM, it is recommended that the GP1 model be recognized as a viable alternative to Akcelik's formula (GL1 model) for green extension time. Although the HCM procedures are typically used to analyze one signalized intersection at a time, the GP1 model was producing superior correlation with CORSIM, even when assuming a uniform flow profile for the single intersection in experiment #1. Ideally, independent researchers could perform independent testing to confirm the results of this paper using different types of input data, e.g. a wider variety of internal link lengths.

Regarding other procedures or applications, the prototype models based on uniform arrivals (GP2, GP3) provide a quick and easy solution for computing actuated phase times. This model typically produces good results when doing calculations by hand, or on a calculator, provided that complexities such as protected-permitted phasing

and spillback are not an issue. In fact, these complexities could be addressed if someone would develop an analytical procedure for queue service time adjustments due to permitted left-turn effects or spillback effects. Still, it would be difficult for analytical procedures to recognize such network-wide complexities as well as simulation can.

Regarding future research, the most pressing need would be to improve the treatment of competing link effects. The queue calibration factor (f_q) formula is able to partially minimize competing link bias. This concept was discussed in chapter 3 and illustrated by the results in chapter 4. The existing formula is only slightly effective. Phase times are significantly biased in the presence of non-critical actuated links with relatively high traffic volumes. An improved queue calibration factor formula, or an improved green extension time calculation based on competing link volumes, would generate higher phase times in the presence of higher competing link volumes. If the competing link volumes were very low, competing link effects would become negligible.

One possible area of future research includes the effects of detector setback on actuated phase times. Detector setback could perhaps be partially handled by computing an effective gap setting [Husch, 1996]. However, the issue is complicated by the fact that TRANSYT-7F queues and flow profiles are tabulated at the stop line of any given link. In order to estimate phase times for a detector far upstream of the stop line, it may be necessary to deduct detector-to-stop-line travel time from queue service times, and to shift the portion of the flow profile being analyzed for computing green extension times.

Another possible area of future research includes the effects of phase skipping on actuated phase times. It is possible that, in the context of the candidate models, reasonable phase times under phase skipping could be obtained by using a minimum

phase time of 0 seconds. Actuated phase times under phase skipping should be tested thoroughly, to determine whether a simple adjustment to the minimum phase time would be effective, or whether adjustments to the models themselves would be necessary.

APPENDIX

Brief Tutorial on Actuated Control

The basic purpose of traffic-actuated control is to reduce vehicle delay through efficient signal timing. It accomplishes this by attempting to terminate the phase as soon as the initial queue of vehicles has been serviced. Phase termination referred to as "gap-out" occurs when a certain sized gap, specified by the gap setting, has been detected in the traffic stream. Presumably, a gap of this size would not normally be observed among the initial queue of vehicles. Phase termination referred to as "max-out" or "force-off" occurs when a pre-specified maximum amount of time elapses without having achieved phase termination via gap-out. Detectors sense the traffic and report vehicle presence to the controller. The controller uses this information in conjunction with the pre-programmed signal settings to produce the signal timing. The size, location, and number of detectors comprise the detector configuration.

Figure A-1 illustrates the impact of signal settings on the phase length. After the signal turns green, the minimum green time must be displayed. After the minimum green has expired, the phase is eligible to be terminated inside the "gap-out region", if detectors

successfully locate a sufficient gap in the traffic stream. If phase termination does not occur in the gap-out region, max-out must occur. Yellow and all red clearance times are displayed following gap-out or max-out. Therefore, because of the gap-out region, the actuated phase length can vary from cycle to cycle. Individual phase times can be as short as (Minimum Green + Y & AR), as long as (Minimum Green + Gap-out Region + Y & AR), or anything in between.

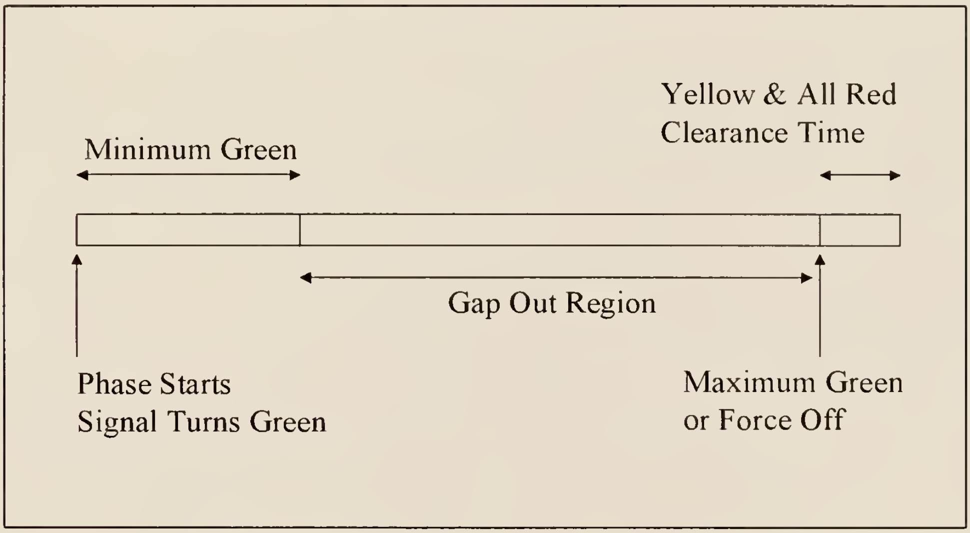


Figure A-1: Actuated Signal Setting Impact on Phase Time

Fully-Actuated Control

When intersections under actuated control are effectively isolated, and not coordinated with other intersections so as to achieve progression, such intersections are often operated under fully-actuated control. In this case, each phase within the signal timing plan is an actuated phase, and subject to termination in response to the traffic flow. In addition, the cycle length is subject to variation in response to the traffic flow. This type of signal control works well when the expected traffic volume demand is fairly

evenly distributed across the intersection approaches. Figure A-2 illustrates a hypothetical intersection under fully-actuated control. It has detectors on each lane, and relatively even volume distribution among the different movements.

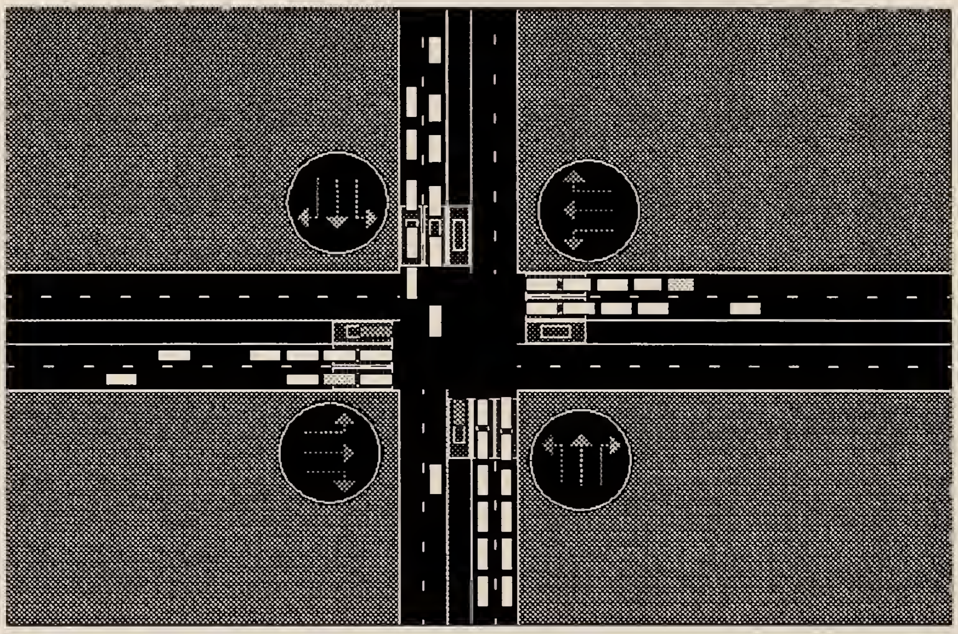


Figure A-2: Typical Conditions Efficiently Handled by Fully-Actuated Control

Semi-Actuated Control, Uncoordinated

When the traffic volume demand is mostly expected along one particular arterial street direction, it may be preferable to allow the signal to rest in green along this heavy volume direction. For example, if the vast majority of traffic is expected in the east-west direction, the east-west phases can be designed as non-actuated, whereas the north-south phases can be designed as actuated. This is called semi-actuated control, in which the minor street receives the minimum amount of green time, and preference is clearly given to the major street. If the major street phase is pre-timed, e.g., always gets 50 seconds of

green time, then the cycle length remains subject to variation due to the minor street actuated phase(s). Alternatively, the major street phase can be non-actuated, and designed to terminate at the "yield point". This allows maintenance of a predictable, background cycle length. Figure A-3 illustrates a hypothetical intersection under semi-actuated control. It has no detectors on the major street through lanes, and predominantly through movement traffic demand.

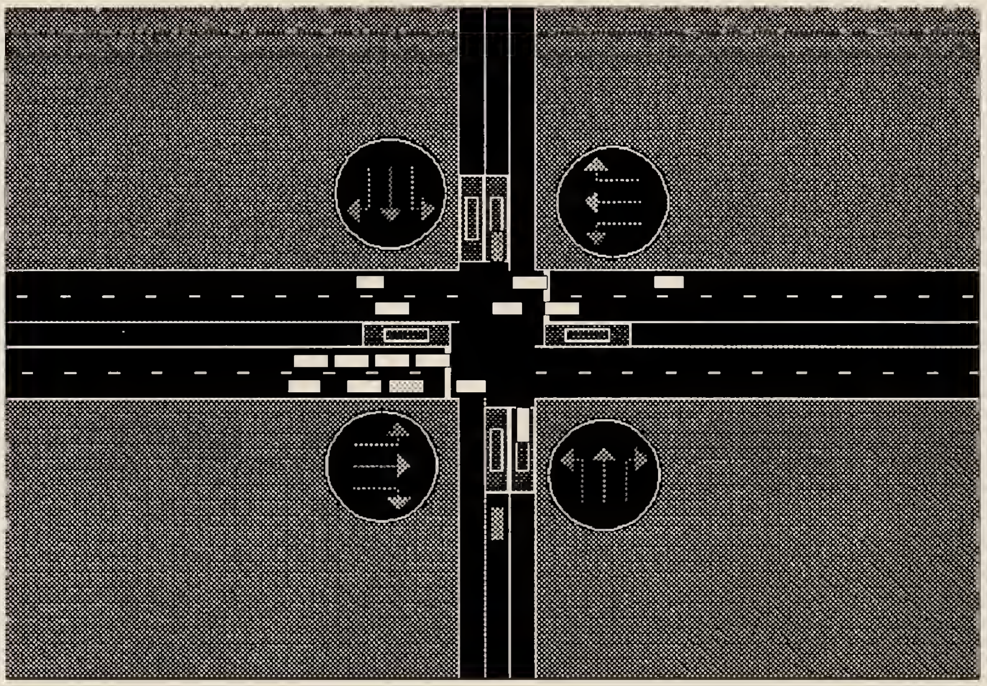


Figure A-3: Typical Conditions Efficiently Handled by Semi-Actuated Control

Semi-Actuated Control, Coordinated

A preferred form of signal control along congested urban arterial streets is semi-actuated control with coordination. This allows for prompt termination of the minor movement phases, and potential early returns to green along the major street. Having a minimum green window along the major street, which terminates reliably at the yield

point, allows for coordination with other signals. The horizontal arrow at the top of Figure A-4 illustrates the minimum green window for the coordinated, non-actuated phase. However, if the actuated minor street phase terminates early, the non-actuated phase may benefit from early return to green. In any case, the minimum green window occurs reliably at the same point in the cycle.

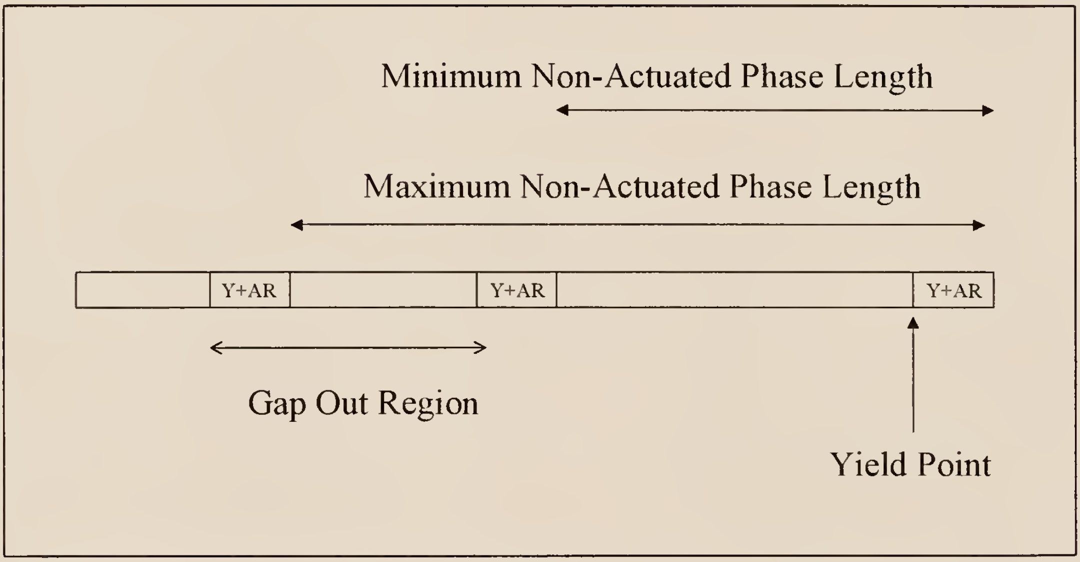


Figure A-4: Actuated Phase Impact on Non-Actuated Phase Time

Brief Tutorial on TRANSYT-7F

TRANSYT-7F is one of the most comprehensive signal timing tools in existence. It is known for its ability to perform network-wide signal timing optimization via iterative simulation runs. It is comprehensive because it has broader capabilities than most signal timing programs. To name just a few, these capabilities include simulation of existing conditions, modeling of a network as opposed to a single arterial or intersection,

separate or simultaneous optimization of splits and offsets, and optimization based on numerous objective functions. With the advent of release 8, it now has the added ability to explicitly model oversaturated traffic conditions. TRANSYT-7F has evolved into a benchmark within the transportation profession. It has facilitated greater understanding of signal timing optimization, while continuing to improve traffic operations as a result of its designs being implemented in the field.

Originally, TRANSYT-7F was developed in an era of pre-timed signal control. Decades ago, the green times at each signal were constant from cycle to cycle, and did not vary in response to the traffic flow. Today, traffic-actuated control has become commonplace in cities throughout the world, taking the place of pre-timed control. In the 1980s, TRANSYT-7F was modified to automatically estimate the average green times for actuated controllers.

The primary purpose of the TRANSYT-7F program is traffic signal timing optimization. However, it is the traffic flow simulation sub-model that gives TRANSYT-7F the unique ability to optimize signal timing effectively under a wide variety of traffic network conditions. In order to understand TRANSYT-7F's strategy for signal timing optimization, it is first necessary to understand the TRANSYT-7F traffic flow sub-model.

Simulation of Traffic Flow

The following description of the traffic flow sub-model appears in the TRANSYT-7F users guide [Wallace et al., 1998]:

> The traffic simulation model in TRANSYT-7F (also called the "traffic model") is among the most realistic of those available in the family of computerized macroscopic traffic models. A macroscopic model is one that considers platoons

of vehicles rather than individual vehicles. TRANSYT-7F simulates traffic flow and small time increments, so its representation of traffic is more detailed than other macroscopic models that assume uniform distributions within the traffic platoons.

The traffic model further utilizes a platoon dispersion algorithm that simulates the normal dispersion (i.e., the "spreading out") of platoons as they travel downstream. It also considers traffic delay, stops, fuel consumption, travel time and other system measures.

Figure A-5 illustrates an example of the platoon dispersion model. In this example, an internal link length of 1000 feet is used. This is also the link length that was used for experiments #2-4 from chapter 4. The flow profile occurring at the downstream stop line is affected by the seven input parameters listed at the top of figure A-5. In addition, multiple upstream links are likely to produce "merging" of such platoons at the downstream stop line.

Because TRANSYT-7F is able to perform mesoscopic simulation in small time increments, it is able to continuously tabulate large quantities of information for each link in the network, and for each step of simulation. For example, during each step of simulation and for each link in the network, TRANSYT-7F keeps track of:

- the arrival rate of traffic volume ("in-pattern")
- departure rate of traffic ("out-pattern")
- saturation flow rate ("go-pattern")
- uniform queue length
- back of queue movement

- link full flag
- and more

Volume (vph)	1000	
Speed (mph)	35	produces the average link travel time (T)
alpha (pdf)	0.35	recommended default for U.S. urban streets
beta	0.8	constant coefficient
Link Length (ft)	1000	
F (smoothing factor)	0.155	based on alpha, beta, and travel time
Travel Time (sec)	19.5	based on speed and link length

	Predicted Flow Rate				
step (t+BT)	v'(t+BT)	F	vt	1-F	v'(t+BT-1)
17	155	0.155	1000	0.845	0
18	286	0.155	1000	0.845	155
19	396	0.155	1000	0.845	286
20	490	0.155	1000	0.845	396
21	569	0.155	1000	0.845	490
22	481	0.155	0	0.845	569
23	406	0.155	0	0.845	481
24	343	0.155	0	0.845	406
25	290	0.155	0	0.845	343
26	245	0.155	0	0.845	290
27	207	0.155	0	0.845	245
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					

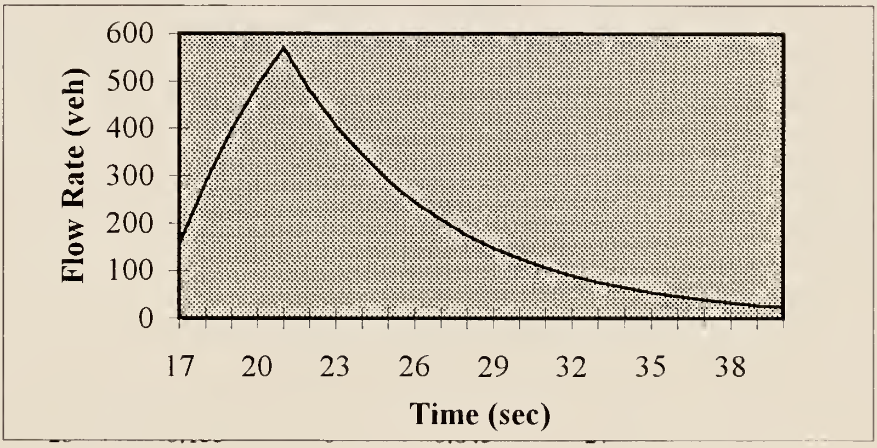


Figure A-5: Typical TRANSYT-7F Platoon Dispersion Model Calculations

Table A-1 shows how the link-specific data is catalogued by TRANSYT-7F during each step of simulation. The program uses a specific two-dimensional array

variable to store each type of data. The first array dimension is for the link number, whereas the second array dimension is for the simulation step number.

Table A-1: TRANSYT-7F Traffic Flow Simulation Arrays

Input Flow Rate (INPAT)						
			Link Number			
			101	102	103	
Simulation		1				
Step Number		2				
		3				

With the advent of step-wise simulation introduced in release 8, the information tabulated by TRANSYT-7F throughout simulation allows modeling of queue spillback and spillover. For each step of simulation that a link is known to be full, saturation flow rates (go-pattern) are reduced accordingly for the upstream feeding links. Information generated from simulation allows the program to produce flow profile diagrams, platoon progression diagrams, time space diagrams, etcetera. Simulation statistics are also used directly to compute output measures of effectiveness including the number of vehicles discharged (“Flow”), link-specific quality of progression (“Arrival Type”), time full%, uniform and random delay, stops, queue, etcetera. Subsequently, these output statistics can be used directly for optimization.

Traffic Signal Timing Optimization

Because detailed output statistics are available for any given simulation run, optimization is theoretically possible by simulating every feasible timing plan and choosing the one that produces desired output. However, there are too many

mathematically feasible timing plans, and not enough time to simulate each one of them. Fortunately, a good solution is usually obtained by performing the hill-climb optimization search technique, in which numerous feasible timing plans are simulated. Mathematics of the hill-climb search technique are described in the users guide, but for now it is only necessary to grasp the concept. The timing plan whose simulation run produces the desired output statistics, according to the user-specified objective function, is reported as optimal. Figure A-6 illustrates the concept graphically.

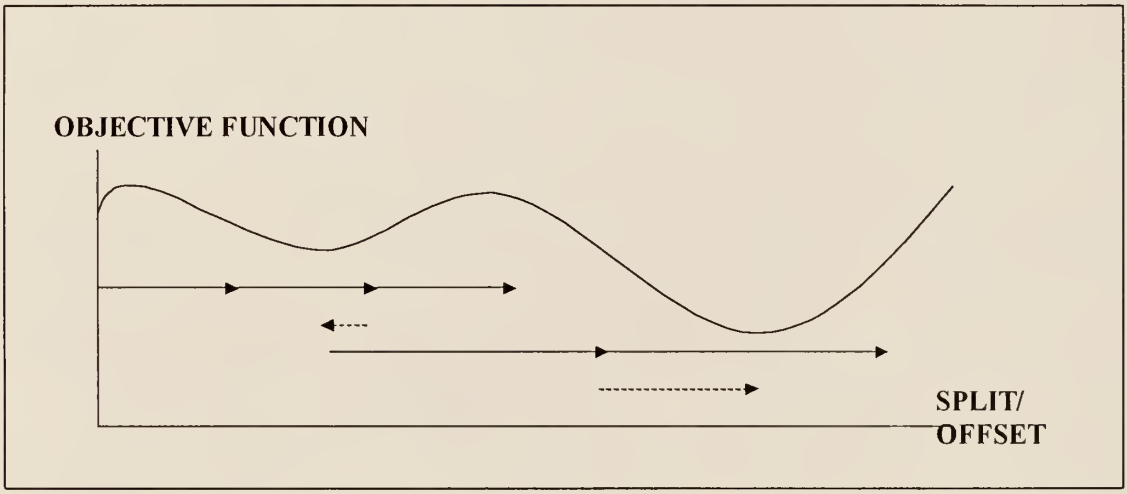


Figure A-6: TRANSYT-7F Hill-Climb Optimization Search Technique

Again quoting from the users guide:

TRANSYT-7F explicitly optimizes phase lengths and offsets for a given cycle length. To determine the best cycle length, an evaluation of a user-specified range of cycle lengths may also be made. To examine alternative phase sequences, multiple computer runs are required. It should be realized that the absolute optimal solution may not be obtained, but TRANSYT-7F has been demonstrated to give reliable signal timings when used with realistic input data.

When optimizing, TRANSYT-7F minimizes (or maximizes, depending on the selection) an objective function called the performance index (PI). The objective function is selected by the user. The best objective function to use depends on the desired operational characteristics of the system under consideration. In many situations, it is a good practice to experiment with the use of different objective functions and select the solution that produces the best results.

Calibration of CORSIM and TRANSYT-7F

In order to test the candidate models, each one was programmed so as to operate in conjunction with TRANSYT-7F. This means that in order to execute one of the candidate models, the updated TRANSYT-7F program was used to process typical TRANSYT-7F input files. Assuming that the candidate models were programmed correctly, the biggest remaining obstacle to compiling useful data and information becomes the differences between TRANSYT-7F and CORSIM themselves.

For example, one possible pitfall would be to observe large differences between results from a candidate model and CORSIM, and conclude the candidate model has poor performance, not realizing that the differences were actually caused by inappropriate differences in the input data for both programs. Keeping the two programs consistent with each other becomes more difficult when they use different input variables to model the same processes.

Queue Discharge Headway (CORSIM), Saturation Flow Rate (TRANSYT-7F)

In experiment #1 for single intersections, all CORSIM queue discharge headways were 2.0 seconds, and all TRANSYT-7F saturation flow rates were 1800 vehicles per

lane per hour (vplph). The average queue discharge headway can be converted to saturation flow rate, and vice versa, with the following formula:

$$\textit{Headway}\ (\text{sec}) = \frac{3600}{\textit{Satflow}\ (\text{vplph})}$$

In experiments #2-4 for arterial streets, TRANSYT-7F saturation flow rates ranged between 1400 and 1900 vplph, based on the benchmark data set [Henry, 1999]. According to the TRANSYT-7F input file format, saturation flow rates are specified on record type 28. Subsequently, these values were directly converted into mean queue discharge headways for usage in the CORSIM input files (record type 11).

Start-up Lost Time

Start-up lost time can be coded into both CORSIM and TRANSYT-7F in units of seconds. In all chapter 4 experiments, start-up lost time was specified as 2 seconds in the TRANSYT-7F input files (record type 1). However, it was not specified as 2 seconds in the CORSIM input files, because CORSIM appears to implement start-up lost time in a different way. The way that CORSIM implements start-up lost time does not appear to be consistent with the HCM or TRANSYT-7F.

For example in CORSIM, when start-up lost time is specified as 2 seconds in the input file, the animation vehicle at the front of a standing queue is noticeably stationary for a full 2 seconds after the signal turns green. This is appropriate. However, it so happens that once a vehicle nudges over the stop line in CORSIM, it is immediately counted as being on the downstream link instead of the upstream link. The initial queued vehicle is essentially discharged instantaneously instead of at the rate of queue discharge headway. In other words, the initial queued vehicle is served in approximately 2 seconds

instead of approximately 4 seconds, i.e. 2 seconds start-up lost time + 2 seconds queue discharge headway. As a result, the effective start-up lost time is something other than 2 seconds.

Figure A-7 illustrates how CORSIM vehicles are no longer considered on the link, once they nudge over the stop line. The gap setting is 3 seconds for the illustrated scenario. In the fourth animation snapshot, the signal turns yellow because no vehicle has been detected for 3 seconds in a row.

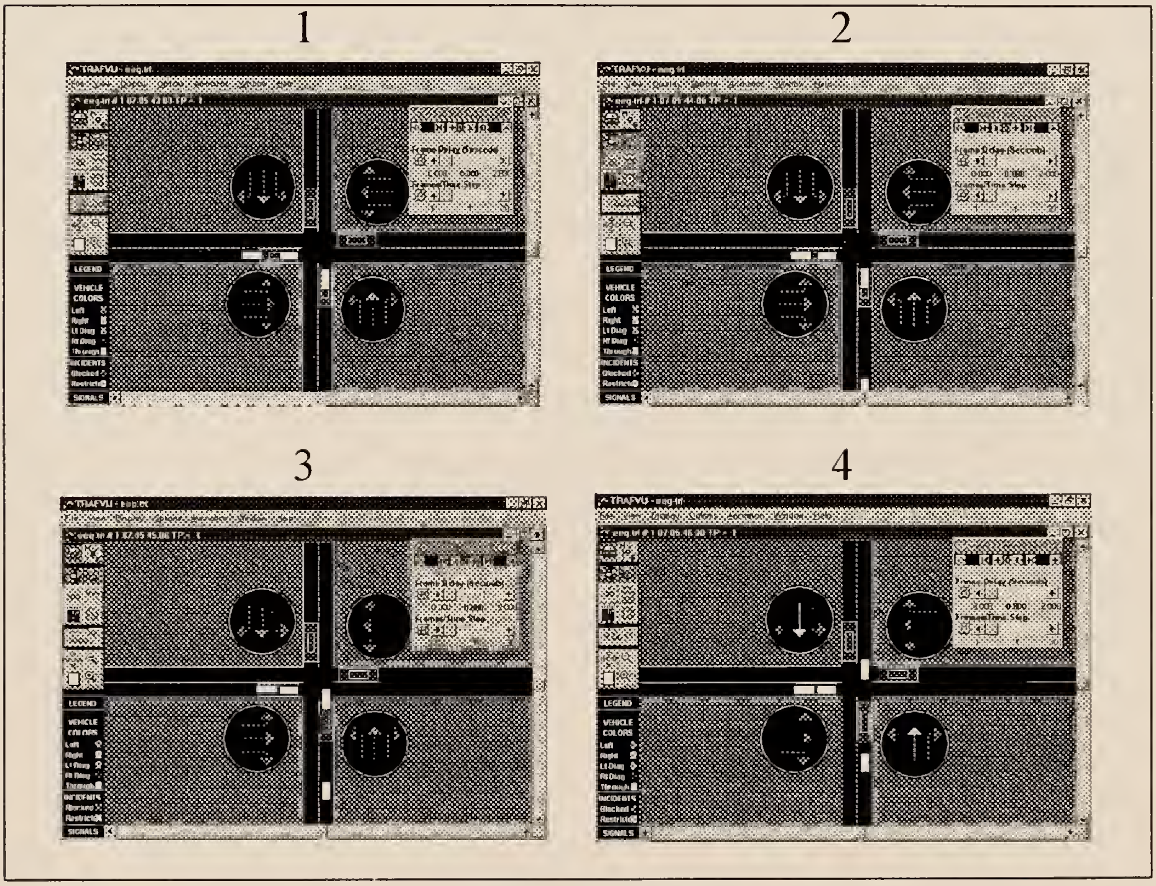


Figure A-7: CORSIM Detector Status Depicting Start-up Lost Time

What start-up lost time should be specified in CORSIM, to be consistent with a 2 second start-up lost time in TRANSYT-7F? The question is complicated by the possible effects of start-up lost time on other vehicles within the queue, besides the lead vehicle. One way to analyze the effects of start-up lost time is to analyze queue service times, which are directly affected by start-up lost time. Ideally, the relationship between queue size and queue service time should be quite predictable. If start-up lost time equals 2 seconds and queue discharge headway equals 2 seconds per vehicle, then the mathematical relationship should be:

$$\text{Queue Service Time} = (\text{Queue Discharge Headway}) \times (\text{Queue Size}) + \text{Startup Lost Time}$$
$$y = 2x + 2$$

For example, a queue of 1 vehicle should be discharged in 4 seconds, a queue of 2 vehicles should be discharged in 6 seconds, etc. Therefore, an experiment was performed in CORSIM using different start-up lost times, in order to determine which start-up lost time would produce queue service times consistent with the equation above. Figure A-8 illustrates some of the results of this test.

Results of the experiment indicated that a start-up lost time of 3.5 seconds produced the closest relationship to $y = 2x + 2$. Start-up lost times of 3 or 4 seconds produced poorer correlation than the results using 3.5 seconds. Therefore, in the chapter 4 experiments, start-up lost time was defined as 3.5 seconds (on record type 11) inside the CORSIM input files.

1	5
2	7
3	8
4	11
4	15
4	8
5	12
5	13
5	10
6	13
6	13
6	15
6	10
6	13
6	12
6	10
6	13
6	14
6	13
6	18
7	15
7	20
7	15
7	15
7	19
7	22
7	15
8	16
8	20
8	18
8	17
8	18
8	16

CORSIM Parameters:

7200	Duration of Simulation (sec)
500	Flow Rate (vph)
5	Intergreen (sec) 5 yellow, 0 all red
3.5	Start-Up Lost Time (sec)

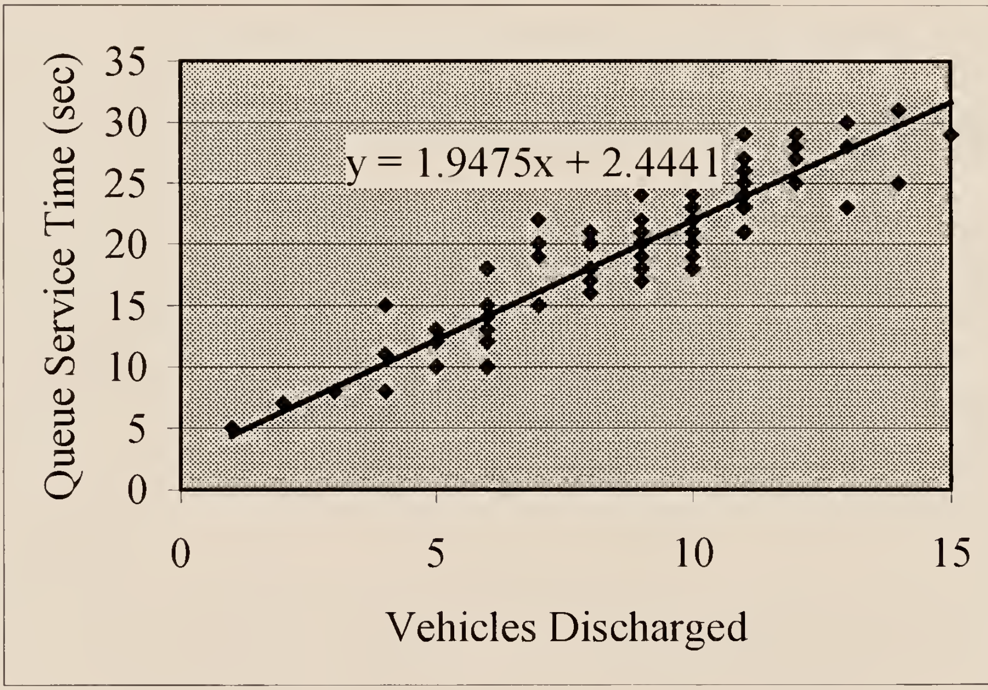


Figure A-8: CORSIM Queue Service Time Experiment Data

Amber Response (CORSIM), Extension of Effective Green (TRANSYT-7F)

TRANSYT-7F and the HCM both implement a parameter called extension of effective green (EEG), a conceptual counterpart to start-up lost time. Whereas start-up lost time temporarily prevents vehicles from moving after the beginning of green, EEG temporarily allows vehicles to continue moving after the end of green. This is consistent with real-world behavior, where drivers will continue to enter the intersection in the

initial seconds of displayed yellow or amber, immediately following green termination. Figure A-9 illustrates that vehicles are considered to be moving after the start-up lost time, and throughout the duration of EEG. Throughout the chapter 4 experiments, EEG was globally defined as 2 seconds within the TRANSYT-7F input files (record type 1).

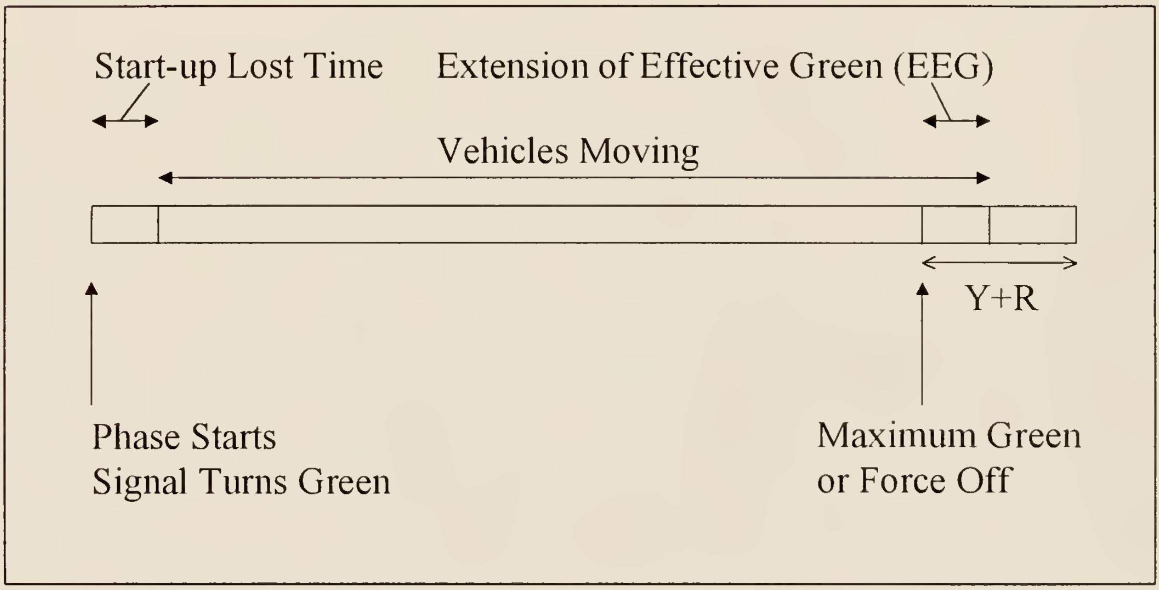


Figure A-9: Graphical Illustration of Extension of Effective Green (EEG)

Alternatively, CORSIM implements a parameter called amber response (record type 144). Based on their individual aggression level, each of the 10 CORSIM driver types possesses a default value for amber response. "The response of drivers to the onset of the amber indication is expressed in terms of an acceptable deceleration" [ITT Systems and Sciences Corp., 1998]. Initial testing was performed to determine the nature of CORSIM flow rate reduction during the yellow interval, based on various user-specified amber response times. This testing showed that it was difficult to come up with amber response input values that would produce identical flow rates to those that would occur

under exactly 2 seconds of EEG. However, testing did show that a yellow time of 2 seconds would produce nearly the maximum flow rate during those 2 seconds along with no flow during the red, i.e. behave very similar to 2 seconds of EEG. Therefore, in order to avoid biases in the results, yellow times in all CORSIM and TRANSYT-7F input files (used in the chapter 4 experiments) were set to exactly 2 seconds.

Internal & External Link Length

Internal link lengths should probably be specified as identical within the input files of CORSIM (record type 11) and TRANSYT-7F (record type 28). They should probably also be specified as identical to those measured in the field as well. This allows for realistic modeling of progression and spillback effects, and similar results between the two programs. Because of differences in platoon dispersion sub-models between the two programs, some differences in progression are currently unavoidable. However, it doesn't seem wise to compensate for these differences by adjusting the internal link lengths. In the chapter 4 experiments #2-4 for arterial streets, internal link lengths were all 1000 feet within the CORSIM and TRANSYT-7F input files.

Figure A-10 illustrates the internal links that separate closely spaced intersections. The dashed lines indicate internal links whose phase times may be susceptible to non-uniform vehicle arrivals or progression effects. On the figure this dashed line looks like perhaps only one link; however, by TRANSYT-7F terminology, the single dashed line contains multiple links because each turning movement is represented by a unique link. Thus, if all three turning movements are being made at both intersections in figure A-10, there are at least six links here, possibly more under complex conditions. By CORSIM

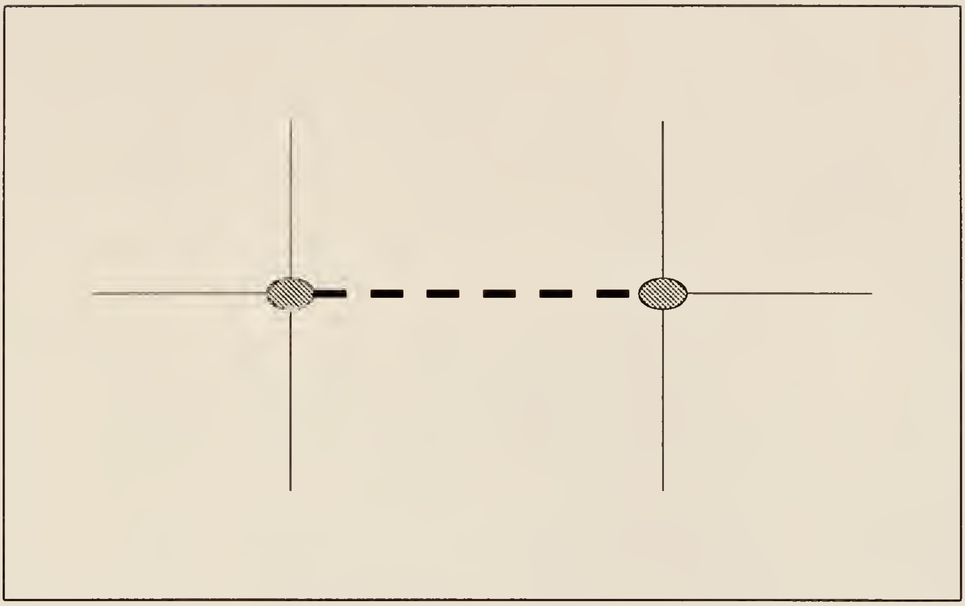


Figure A-10: Internal Links

terminology, the single dashed line contains two uni-directional links, e.g. (1-2) and (2-1) if the nodes were numbered 1 and 2. The regular lines indicate external links where vehicle arrivals are typically assumed to be uniform or random.

External links can be thought of as coming out of a shopping mall, or residential area, or any area with no nearby upstream signal. This means that vehicle arrivals may also be assumed uniform or random, and links assumed external, when the distance between traffic signals becomes sufficiently large such that platoons of vehicles have enough time and space to spread out. It is unclear exactly what distance is sufficient to dissipate the platoons and render progression effects negligible. This distance is hypothesized to be somewhere between 1 and 3 miles.

Although it seems clear that internal link lengths should be specified as identical within the input files of CORSIM and TRANSYT-7F, and identical to those measured in

the field as well, the appropriate coding of external link lengths is less clear. External link lengths should perhaps be a function of the type of analysis being conducted. It can have an effect on outputs related to queue length, travel time, etc. There is no recommended default value in the literature, nor should there be. Some of the strategies used by practitioners for external link length coding are as follows.

High priority links

If the external link serves a specific facility, e.g. shopping mall, church, residential area or school, then queue spillback into these areas can be a concern. In these cases, the external link length is coded such that if the link becomes full, spillback is deemed unacceptable. In these situations, the specific external link length to be coded is absolutely a function of engineering judgment. For example, if only 200 feet of roadway separate a signal from a stop-controlled intersection serving a shopping mall, then perhaps the external link should be coded as 200 feet in order to analyze whether spillback problems will tie up circulating mall traffic. In CORSIM animation, spillback problems become apparent when the queue visibly reaches the entry node. In TRANSYT-7F, optimization objective functions can be customized to produce designs that improve external link performance.

Low priority links

If an engineer is not concerned about the performance of a certain external link, they will sometimes specify a very short link length in CORSIM or TRANSYT-7F. The potential advantages of this coding strategy are threefold. First, program running time is reduced. Second, network-wide output statistics better reflect performance on links that the engineer cares about. Short external link lengths do not compromise accuracy of the

internal link results. The flow rate of traffic coming into the network from the external link is unaffected by its link length. Third, when TRANSYT-7F sees a very short link length, the effect of this link's performance on the optimization objective function is automatically diminished. The TRANSYT-7F input format is such that external links coded with a length of 0 are automatically assigned a token link length of 100-200 feet for simulation purposes, but are effectively removed from the link list for optimization purposes.

Medium priority links

If an external link is deemed to have medium priority, the link length is sometimes coded as a medium-large distance such as one half-mile (2640 feet). Because the link length is not too large, network-wide results are not inappropriately biased, and program running time (CORSIM only) is not unnecessarily increased. Because the link length is not too short, queuing activity can be fully analyzed, and link performance affects the optimization objective function (TRANSYT-7F only).

External link lengths for chapter 4 experiments

In the chapter 4 experiments, external link lengths were coded for both programs such that queuing activity could be fully analyzed on undersaturated actuated phases. On oversaturated actuated phases, external link length is not important because it is clear that the phase will terminate via max-out each time. However, on undersaturated actuated phases, the queue length directly affects the phase length. In order to correctly simulate queue service time, CORSIM and TRANSYT-7F external links must be long enough to hold the entire undersaturated queue. Therefore, external link lengths for the chapter 4

experiments were coded to be long enough to hold any undersaturated queue, but short enough so that optimization experiment #4 results would not be compromised.

Free Flow Speed

Free flow speed is specified in units of miles per hour in both CORSIM and TRANSYT-7F. This parameter affects platoon dispersion, progression, and actuated phase green extension time. Free flow speeds were coded with the same values for each program (CORSIM record type 11, TRANSYT-7F record type 28). In experiment #1 from chapter 4, free flow speeds were 25 mph. In experiments #2-4, free flow speeds were 30 mph on external links and 35 mph on internal links.

Lane Usage, Channelization, or Designation

In TRANSYT-7F, lane channelization is handled implicitly by a couple of individual inputs. Shared lane links are listed on record type 7. Shared lanes that are served by protected-only phasing are often specified as a single link, instead of as multiple links in a shared lane. Referring to figure A-11, the protected shared right can be defined with a single link number (e.g. 104). Proper link definitions on record type 28 are important. For example, if a shared left-turn lane served by permitted phasing sits adjacent to an exclusive through lane, three separate links should be defined to reflect complex performance. The first link represents the permitted left-turn, the second link represents the shared through lane affected by the permitted left, and the third link represents the exclusive through lane unaffected by the permitted left. Referring to figure A-11, the permitted shared left should be defined with two link numbers (e.g. 101 & 102). Links 101 and 102 should also be listed as shared lane links on record type 7.

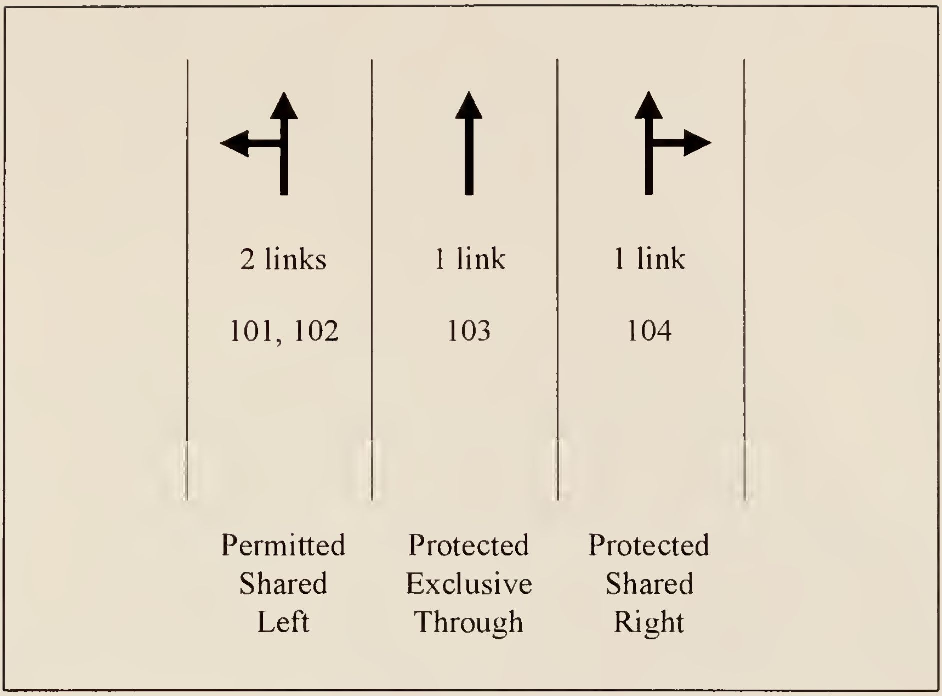


Figure A-11: TRANSYT-7F Lane Channelization Coding Technique

In the chapter 4 experiments, shared right-turn lanes were always served by protected-only phasing, and were thus defined using a single link number. Shared left-turn lanes in the chapter 4 experiment data were defined using two link numbers; however, permitted shared lefts were not analyzed. This should not significantly compromise the chapter 4 results, because actuated-permitted-shared lefts are relatively rare.

In CORSIM, lane channelization is specified explicitly by lane-specific channelization codes on record type 11. The appropriate channelization codes in CORSIM are independent of signal phasing. Figure A-12 illustrates sample CORSIM channelization codes. Such codes were used as appropriate in the chapter 4 experiment data. The figure also shows that the three lanes are defined as 1 uni-directional link, although each lane on the link requires a channelization code. This contrasts with

TRANSYT-7F, in which the three lanes are defined as 4 links without channelization codes.

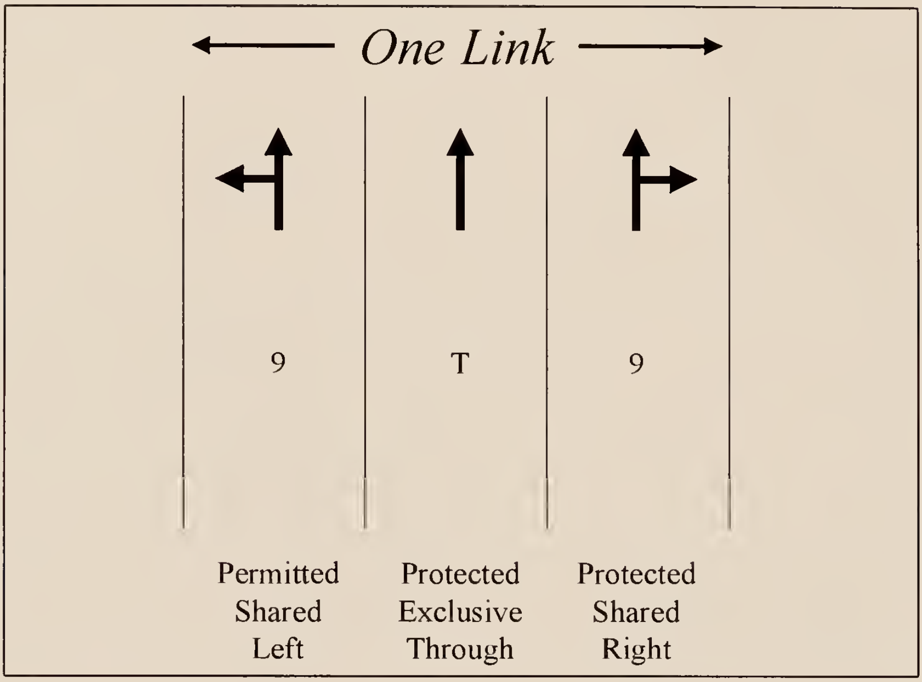


Figure A-12: CORSIM Lane Channelization Coding Technique

Entry Node and Approach Volumes (CORSIM), Link Volumes (TRANSYT-7F)

In CORSIM, vehicles are introduced into the network by the entry nodes to external links. At each entry node, the total hourly traffic volume entering the network is specified on record type 50. Exiting traffic volume does not need to be specified. For the internal nodes that represent intersections, turning movement percentages are specified on record type 21. Turning movement volumes may optionally be specified here, but CORSIM subsequently converts these into percentages for use internally. In the chapter 4 experiment data, entry node volumes and turning movement volumes were specified as appropriate.

In TRANSYT-7F, hourly volumes for each link are specified on record type 28. Summation of external link volumes should be equal to the entry node volumes from CORSIM. Turning movement percentages or volumes are handled automatically by the link volumes in conjunction with feeding link volumes that are also specified on record type 28. TRANSYT-7F uses a traffic flow balancing algorithm to handle unequal input volumes on upstream-downstream links. The program issues a fatal error message if the link volume discrepancy becomes too large. In experiment #1 from chapter 4, link volumes were always 300, 500, or 700 vehicles per hour. In experiments #2-3, link volumes were taken directly from the benchmark data sets [Henry, 1999]. In experiment #4, benchmark data set volumes were doubled (using record type 36) on the entire network in order to create oversaturation and spillback conditions. Volume multiplier record type 36 allows a volume adjustment factor to be applied globally, or to a specific list of links.

Automated Process: Experimental Version of TRANSYT-7F

The T7FACT program was developed in order to implement the candidate models. It essentially runs TRANSYT-7F iteratively; however, it also reads flow profile and queue length data from the debug file, performs the calculations necessary to develop a new timing plan, and tests for convergence. By default, the process terminates whenever it observes that the derived signal timing plan is identical to that of any previous iteration. Alternatively, the user can specify a desired number of model iterations. Figure A-13 illustrates the automated process.

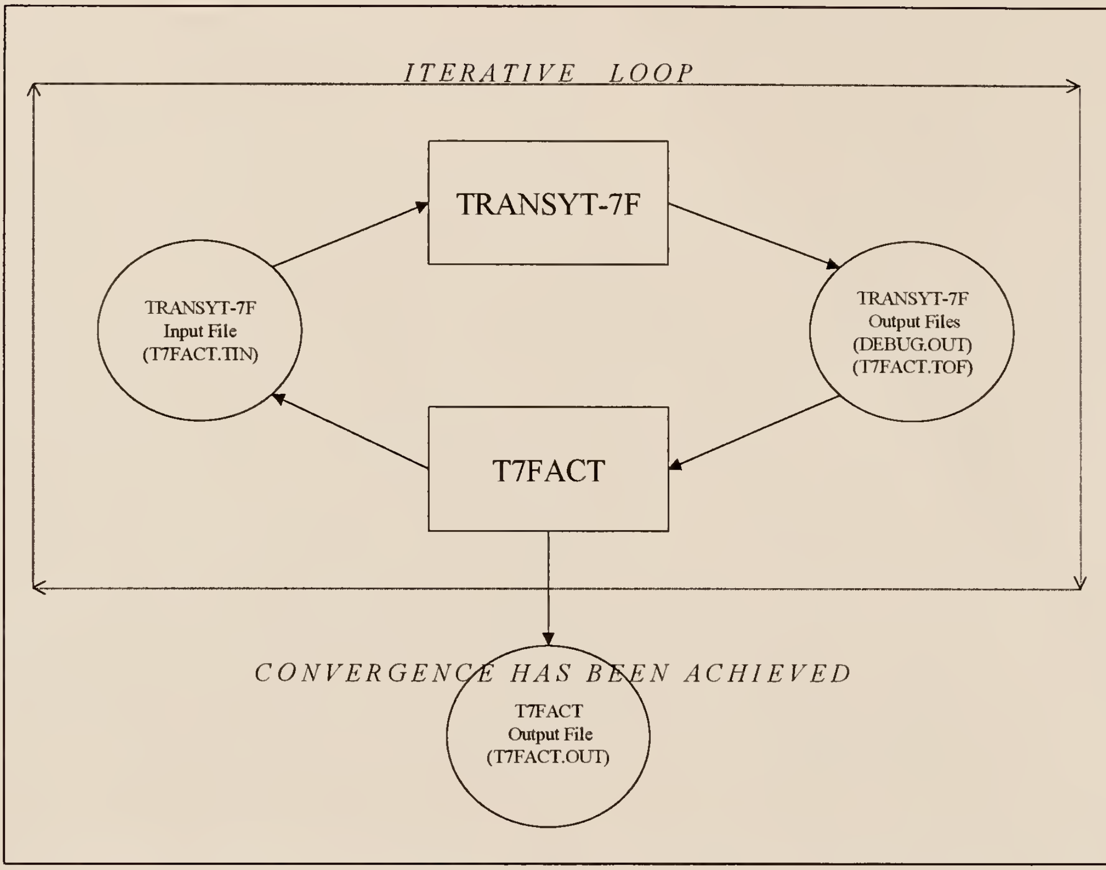


Figure A-13: Automated Process Controlled by T7FACT

The first operation performed by T7FACT is to request an initial TRANSYT-7F run. The results of the first-iteration run are used to determine whether the existing phase lengths accurately represent actuated control operation. If not, additional iterations will be needed in order to determine the appropriate phase lengths. Before making any runs, it is necessary to create an input file that specifies the modeling conditions and criteria. It is important to note that any user-specified signal timing plan within the standard TRANSYT-7F input file is interpreted by T7FACT as force-offs, or maximum green settings, and the candidate models use these directly as the initial timing plan in the iterative phase time calculation process. If the user does not specify an initial timing plan

within the TRANSYT-7F input file, a preliminary run must be performed in order to obtain an initial timing plan from TRANSYT-7F.

After TRANSYT-7F has terminated, T7FACT retrieves information from two of its output files. It reads flow profile and queue length data from the debug output file in order to develop a new timing plan, and then test for convergence. If convergence has not been achieved, the new design is exported into the TRANSYT-7F input file, and the iterative process continues. It also post-processes the overall average delay per vehicle that is listed in the standard TRANSYT-7F output file.

The T7FACT output file is generated to illustrate the results of candidate model execution. Its outputs include:

- user-selected queue service time and green-extension time models
- queue service time (g_s) and green extension time (g_e)
- minimum and maximum green time; minimum and maximum phase time
- model-estimated phase time
- offset value and offset reference phase
- queue calibration factor (f_q)
- intermediate and final signal timing plans
- intermediate and final average vehicle delay reported by TRANSYT-7F

REFERENCES

Akcelik, R., "Analysis of Vehicle-Actuated Signal Operations," Australian Road Research Board, Working Paper WD TE 93/007, 1993.

Akcelik, R., "Estimation of Green Times and Cycle Time for Vehicle-Actuated Signals," Transportation Research Record 1457, TRB, Washington D.C., 1994.

Akcelik, R. and M. Besley, SIDRA 5 User Guide, ARRB Transport Research Ltd., Vermont South, Victoria, Australia, January, 1999.

Bonneson, J. A. and P. T. McCoy, "A Methodology for Evaluating Traffic Detector Designs," TRB 72nd Annual Meeting, Washington D.C., January 1993.

Courage, K.G., Akcelik, R., and D. Fambro, "Capacity Analysis of Traffic-Actuated Intersections, NCHRP Project 3-48," prepared for NCHRP by the University of Florida Transportation Research Center, Texas Transportation Institute, and ARRB Transport Research Ltd., December 1996.

Courage, K. G. and C. E. Wallace. Methodology for Optimizing Signal Timing: MOST Volume 1. Gainesville, Florida, December 1991.

Courage, K. G. and J. Luh. Development of Guidelines for Implementing Computerized Timing Designs at Traffic Actuated Signals. Gainesville, Florida, December 1989.

Cribbins, P. E. and C. A. Meyer, "Choosing Detectors and Placing Them Right," Traffic Engineering, Vol. 45, No. 1, January 1975, pp. 15-18.

Deakin, E.A., May, A.D. and A. Skabardonis, "Energy Savings from Signal Timing Optimization: Evaluation of California's Statewide Program," Compendium of Technical Papers, Institute of Transportation Engineers, San Francisco, CA, 1984.

Fambro, D. B., Sunkari, S. R., Hubbard, S. M., and Irvine, Y. D. "Implementation Guidelines for Retiming Isolated Intersections," Report No. FHWA/TX-93/1164-1, Texas Transportation Institute, December 1992.

Halati, A. and J. F. Torres. "Enhancement of the Value Iteration Program Actuated Signals - Part 2," PennDOT Research Project Number 88-12, Pacific Palisades, CA., February 1992.

Hale, David, "How Many Netsim Runs Are Enough?" McTrans Volume 11 Number 3, Gainesville, FL, March 1997.

Hale, David, "Design and Optimization Strategies for Traffic Actuated Signal Timing Parameters", Masters Degree Report, Gainesville, FL, December 1995.

Henry, Micah A., "The Generation of an Arterial Benchmark Data Set and Its Application to the Effect of Phase Sequence on Arterial Progression," Masters Degree Report, Gainesville, FL, 1999.

Husch, David, "The Percentile Delay Method," Trafficware Internet Technical Support Articles, Albany, CA, August 1996.

________, Corsim Users Manual Version 1.04, ITT Systems and Sciences Corporation, Colorado Springs, CO, March, 1998.

Kell, J. H. and T. J. Fullerton, Manual of Traffic Signal Design 2nd Edition, Institute of Transportation Engineers, Prentice-Hall, Englewood Cliffs, NJ, 1991.

Khatib, Z. K. and P. B. Coffelt, "Conversion of Signal Timing Plans into Traffic Controller Parameters," ITE Journal, Moscow, ID, August, 1999.

Krzeminski, James, "Analysis of Surveillance Data from Coordinated Arterial Traffic Control Systems," Masters Degree Report, Gainesville, FL, December 1996.

Lin, Feng-Bor, "Optimal Timing Setting and Detector Lengths of Presence Mode Full-Actuated Control," Transportation Research Record 1010, TRB, National Research Council, Washington D.C., 1985.

Lin, Pei-Sung, "Capacity Analysis for Traffic Actuated Intersections Draft Literature Review," Working Paper NCHRP 3-48-2, Gainesville, FL, May 1994.

Lomax, T. and D. Schrank, "Traffic Jams: Not Just Big-City Problem," Texas Transportation Researcher Vol. 34 No. 3, College Station, TX, 1998.

Orcutt, F.L., The Traffic Signal Book, Arlington, TX, 1993.

Robertson, D.I., "TRANSYT: Traffic Network Study Tool," Fourth International Symposium on the Theory of Traffic Flow, Karlsruhe, Germany, 1968.

Skabardonis, Alexander, Progression Through a Series of Intersections with Traffic Actuated Controllers. Volume 2, User's Guide, Report No. FHWA-RD-89-133, October 1988.

Skabardonis, A. and A. Weinstein, "TRANSYT-7FC, TRANSYT Model for Actuated Signals," Institute for Transportation Studies, University of California, Berkeley, CA, April 1988.

Tarnoff, P. and P. S. Parsonson, "Selecting Traffic Signal Control at Individual Intersections," NCHRP report 233, Transportation Research Board, National Research Council, Washington D.C., June 1981.

________, Highway Capacity Manual, Special Report 209, Transportation Research Board, 1985, 1994, 1997.

Wallace, C.E., K.G. Courage and D.P. Reaves, "National Signal Timing Optimization Project -- Final Evaluation Report," prepared for the Federal Highway Administration by the University of Florida Transportation Research Center, 1981.

Wallace, C.E., K.G. Courage, M.A. Hadi, and A.C. Gan, TRANSYT-7F Users Guide, Volume 4 of a series prepared for FHWA by COURAGE AND WALLACE, Gainesville, FL, March 1998.

BIOGRAPHICAL SKETCH

David Kivilcim Hale was born in Philadelphia, PA, and raised in the town of Norwich, VT, near Hanover, NH, and Dartmouth College. In the 1970s and early 80s his parents both worked at Dartmouth. In 1983 the family moved to Tucson, AZ, after Frank Hale earned a faculty position at the University of Arizona's College of Medicine. David graduated from Tucson's Amphitheater High School in May 1988, and earned a civil engineering bachelor's degree from the University of Arizona in May 1993.

Since then David has pursued graduate degrees in transportation engineering at the University of Florida, while providing software testing and technical support for the McTrans center. He earned a Master of Engineering degree in December 1995 before enrolling in the Ph.D. program. He has also participated in the Institute of Transportation Engineers (ITE) as a member of the University of Florida student chapter and the Florida section.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Kenneth G. Courage, Chairman
Professor of Civil and
Coastal Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Joseph P. Geunes
Assistant Professor of Industrial and
Systems Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Scott S. Washburn
Assistant Professor of Civil and
Coastal Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Charles E. Wallace
Engineer of Civil and
Coastal Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William M. Sampson
Associate in Engineering of Civil
and Coastal Engineering

This thesis was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

M. J. Ohanian
Dean, College of Engineering

Winfred M. Phillips
Dean, Graduate School